

# **A genome-wide association meta-analysis of self-reported allergy identifies shared and allergy-specific susceptibility loci**

## **Supplementary Tables, Figures and Note**

David A. Hinds<sup>1</sup>, George McMahon<sup>2</sup>, Amy K. Kiefer<sup>1</sup>, Chuong B. Do<sup>1</sup>, Nicholas Eriksson<sup>1</sup>, David M. Evans<sup>2</sup>, Beate St Pourcain<sup>3</sup>, Susan M. Ring<sup>3</sup>, Joanna L. Mountain<sup>1</sup>, Uta Francke<sup>1</sup>, George Davey-Smith<sup>2</sup>, Nicholas J. Timpson<sup>2\*</sup>, Joyce Y. Tung<sup>1\*</sup>

<sup>1</sup> 23andMe, Inc., Mountain View, CA 94043, USA

<sup>2</sup> Medical Research Council (MRC) Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom

<sup>3</sup> School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom

\* Equal contributions

Supplementary Table 2. Detailed results for allergy loci in the 23andMe cohort.

<b>SNP</b>	<b>region</b>	<b>position</b>	<b>alleles</b>	<b>RAF</b>	<b>r<sup>2</sup></b>	<b>typed</b>	<b>OR</b>	<b>95% CI</b>	<b>P<sub>main</sub></b>	<b>P<sub>int</sub></b>	<b>P<sub>total</sub></b>
rs2101521	4p14	38811551	A/G	0.762	0.982	I	1.138	[1.105,1.172]	2.4e-17	0.00015	1.5e-19
rs1438673	5q22.1	110467499	T/C	0.495	0.988	I	1.116	[1.089,1.144]	2.3e-18	0.0062	3.6e-19
rs2155219	11q13.5	76299194	G/T	0.515	0.990	P	1.108	[1.081,1.136]	2.5e-16	0.63	3.2e-15
rs10189629	2q12.1	102879464	A/C	0.855	0.995	P	1.152	[1.111,1.194]	1.2e-14	0.058	2.0e-14
rs6906021	6p21.32	32626311	T/C	0.475	0.951	P	1.097	[1.070,1.125]	3.4e-13	0.10	7.4e-13
rs9266772	6p21.33	31352113	T/C	0.188	0.995	P	1.101	[1.068,1.135]	9.2e-10	0.79	1.4e-08
rs7720838	5p13.1	40486896	G/T	0.575	0.997	G	1.086	[1.060,1.113]	6.6e-11	0.060	9.2e-11
rs10497813	2q33.1	198914072	T/G	0.482	0.999	I	1.075	[1.050,1.102]	5.5e-09	0.54	5.7e-08
rs9860547	3q28	188128979	G/A	0.463	0.948	I	1.073	[1.047,1.100]	2.8e-08	0.14	7.9e-08
rs7032572	9p24.1	6172380	A/G	0.166	0.668	I	1.129	[1.085,1.175]	2.0e-09	0.95	3.6e-08
rs6021270	20q13.2	50141264	C/T	0.942	0.968	I	1.173	[1.111,1.238]	9.5e-09	0.12	2.8e-08
rs9303280	17q12	38074031	T/C	0.520	0.980	I	1.067	[1.041,1.094]	2.0e-07	0.059	2.3e-07
rs17228058	15q22.33	67450305	A/G	0.235	0.991	I	1.084	[1.053,1.116]	3.5e-08	0.45	2.8e-07
rs962993	10p14	9053132	T/C	0.571	0.931	I	1.073	[1.045,1.100]	7.9e-08	0.54	7.9e-07
rs17388568	4q27	123329362	G/A	0.275	1.000	G	1.074	[1.045,1.103]	2.9e-07	0.040	2.1e-07
rs1998359	14q21.1	38077148	C/G	0.249	0.839	I	1.080	[1.047,1.113]	9.6e-07	0.86	1.3e-05
rs6473223	8q21.13	81268155	C/T	0.360	0.980	I	1.061	[1.034,1.088]	6.7e-06	0.18	2.1e-05
rs10174949	2p25.1	8442248	A/G	0.727	1.000	G	1.070	[1.041,1.099]	1.3e-06	0.016	4.0e-07
rs7203459	16p13.13	11230703	C/T	0.733	0.989	P	1.074	[1.044,1.104]	5.0e-07	0.020	2.3e-07
rs2107357	16p12.1	27410829	G/A	0.138	0.967	I	1.091	[1.054,1.131]	1.2e-06	0.63	1.2e-05
rs2056417	1p36.22	10581658	A/G	0.697	1.000	G	1.058	[1.030,1.086]	3.2e-05	0.55	0.00025
rs10893845	11q24.3	128186882	T/G	0.493	0.999	G	1.063	[1.038,1.089]	8.2e-07	0.44	6.2e-06
rs17533090	6p21.32	32590722	G/T	0.840	0.998	G	1.130	[1.093,1.169]	9.6e-13	8.4e-13	2.1e-22

Region, cytogenetic band; position, build 37 map position of SNP; alleles, low/high risk alleles on genomic reference strand; RAF, risk allele frequency; r<sup>2</sup>, imputation quality estimated from variance in imputed allele dosages; typed: I=imputed, G=genotyped, P=partial (genotyped in ~65% of individuals); OR, odds ratio per copy of the risk allele, for the main effect across allergy types; CI, confidence interval; P<sub>main</sub>, test of main effect across allergens; P<sub>int</sub>, test for interaction with allergy type; P<sub>total</sub>, joint test including main effect and interaction.

Supplementary Table 3. Detailed results for allergy loci in the ALSPAC cohort.

<b>SNP</b>	<b>region</b>	<b>position</b>	<b>alleles</b>	<b>RAF</b>	<b>r<sup>2</sup></b>	<b>typed</b>	<b>OR</b>	<b>95% CI</b>	<b>P<sub>main</sub></b>	<b>P<sub>int</sub></b>	<b>P<sub>total</sub></b>
rs2101521	4p14	38811551	A/G	0.791	0.956	I	1.262	[1.139,1.398]	8.2e-06	0.75	0.00013
rs1438673	5q22.1	110467499	T/C	0.518	0.995	I	1.124	[1.041,1.214]	0.0028	0.60	0.019
rs2155219	11q13.5	76299194	G/T	0.489	0.946	I	1.177	[1.089,1.273]	4.2e-05	0.16	0.00015
rs10189629	2q12.1	102879464	A/C	0.870	0.997	I	1.200	[1.062,1.355]	0.0033	0.90	0.030
rs6906021	6p21.32	32626311	T/C	0.475	0.982	I	1.117	[1.032,1.208]	0.0060	0.56	0.031
rs9266772	6p21.33	31352113	T/C	0.222	0.995	I	1.188	[1.084,1.303]	0.00023	0.034	0.00011
rs7720838	5p13.1	40486896	G/T	0.616	0.999	G	1.035	[0.954,1.123]	0.41	0.29	0.37
rs10497813	2q33.1	198914072	T/G	0.489	0.999	I	1.084	[1.002,1.172]	0.043	0.97	0.24
rs9860547	3q28	188128979	G/A	0.455	0.947	I	1.118	[1.030,1.213]	0.0077	0.44	0.032
rs7032572	9p24.1	6172380	A/G	0.171	0.975	I	1.063	[0.960,1.177]	0.24	0.40	0.35
rs6021270	20q13.2	50141264	C/T	0.918	0.945	I	1.094	[0.937,1.277]	0.25	0.41	0.38
rs9303280	17q12	38074031	T/C	0.501	0.986	I	1.111	[1.028,1.200]	0.0082	0.54	0.040
rs17228058	15q22.33	67450305	A/G	0.269	0.971	I	1.065	[0.977,1.161]	0.15	0.71	0.42
rs962993	10p14	9053132	T/C	0.612	0.893	I	1.082	[0.994,1.178]	0.069	0.68	0.25
rs17388568	4q27	123329362	G/A	0.272	0.996	G	1.095	[1.004,1.193]	0.040	0.46	0.12
rs1998359	14q21.1	38077148	C/G	0.229	0.928	I	1.128	[1.029,1.236]	0.010	0.60	0.053
rs6473223	8q21.13	81268155	C/T	0.346	0.995	I	1.156	[1.067,1.252]	0.00041	0.77	0.0043
rs10174949	2p25.1	8442248	A/G	0.706	1.000	G	1.101	[1.013,1.198]	0.024	0.99	0.16
rs7203459	16p13.13	11230703	C/T	0.741	0.969	I	1.065	[0.971,1.168]	0.18	0.25	0.21
rs2107357	16p12.1	27410829	G/A	0.136	0.961	I	1.096	[0.978,1.227]	0.11	0.83	0.40
rs2056417	1p36.22	10581658	A/G	0.671	1.000	G	1.175	[1.080,1.278]	0.00017	0.89	0.0022
rs10893845	11q24.3	128186882	T/G	0.493	0.997	G	1.038	[0.961,1.121]	0.35	0.99	0.82
rs17533090	6p21.32	32590722	G/T	0.830	0.897	G	1.104	[0.989,1.232]	0.077	0.063	0.038

Region, cytogenetic band; position, build 37 map position of SNP; alleles, low/high risk alleles on genomic reference strand; RAF, risk allele frequency; r<sup>2</sup>, imputation quality estimated from variance in imputed allele dosages; typed: I=imputed, G=genotyped; OR, odds ratio per copy of the risk allele for the main effect across allergy types; CI, confidence interval; P<sub>main</sub>, P value for test of main effect across allergy types; P<sub>int</sub>, test for interaction with allergy type; P<sub>total</sub>, overall test including main effect and interaction.

Supplementary Table 4. Tests for platform effects, age effects, and study heterogeneity

SNP	$P_{gt}$	$P_{age}$	$P_Q$	$I^2$	95% CI	gene context
rs2101521	0.46	0.057	0.06	0.72	[0.00,0.92]	TLR1-[]--TLR6
rs1438673	0.33	0.57	0.86	0.00	[0.00,0.80]	WDR36-[]--CAMK4
rs2155219	0.69	0.56	0.15	0.53	[0.00,0.89]	C11orf30-[]--LRRC32
rs10189629	0.15	0.16	0.53	0.00	[0.00,0.80]	IL1RL2-[]--IL1RL1
rs6906021	0.17	0.41	0.68	0.00	[0.00,0.80]	HLA-DQA1-[]--HLA-DQB1
rs9266772	0.64	0.98	0.12	0.58	[0.00,0.90]	HLA-C---[]--MICA
rs7720838	0.91	0.89	0.27	0.19	[0.00,0.84]	[]--PTGER4
rs10497813	0.96	0.83	0.85	0.00	[0.00,0.80]	[PLCL1]
rs9860547	0.31	0.56	0.35	0.00	[0.00,0.80]	[LPP]
rs7032572	0.098	0.99	0.28	0.15	[0.00,0.83]	RANBP6---[]--IL33
rs6021270	0.39	0.24	0.41	0.00	[0.00,0.80]	[NFATC2]
rs9303280	0.67	0.59	0.34	0.00	[0.00,0.80]	[GSDMB]
rs17228058	0.94	0.37	0.70	0.00	[0.00,0.80]	[SMAD3]
rs962993	0.34	0.66	0.85	0.00	[0.00,0.80]	GATA3---[]
rs17388568	0.26	0.052	0.68	0.00	[0.00,0.80]	[ADAD1]
rs1998359	0.96	0.79	0.37	0.00	[0.00,0.80]	FOXA1-[]---TTC6
rs6473223	0.43	0.15	0.05	0.75	[0.00,0.93]	TPD52---[]---ZBTB10
rs10174949	0.20	0.42	0.52	0.00	[0.00,0.80]	[]--ID2
rs7203459	0.26	0.27	0.86	0.00	[0.00,0.80]	[CLEC16A]
rs2107357	0.041	0.67	0.95	0.00	[0.00,0.80]	IL4R-[]--IL21R
rs2056417	0.24	0.42	0.02	0.82	[0.00,0.94]	[PEX14]
rs10893845	0.51	0.40	0.56	0.00	[0.00,0.80]	[]--ETS1

$P_{gt}$ , result of test for a genotyping platform effect between Illumina HumanHap 550-based results and OmniExpress-based results in the 23andMe cohort;  $P_{age}$ , test of interaction between genotype and age in the 23andMe cohort;  $P_Q$ ,  $P$  value for test of heterogeneity based on Cochran's Q statistic;  $I^2$ , proportion of variance in effects attributable to study heterogeneity; 95% CI, confidence interval for  $I^2$ ; gene context, as in Table 2.

Supplementary Table 6. Replication of allergy loci in analysis of allergic sensitization<sup>20</sup>.

SNP	OR	95% CI	match	P	P <sub>fisher</sub>	gene context
rs2101521	1.234	[1.155,1.319]	Yes	6.3e-10	3.7e-30	TLR1-[]--TLR6
rs1438673	1.114	[1.060,1.170]	Yes	2.2e-05	5.0e-25	WDR36-[]--CAMK4
rs2155219	1.202	[1.142,1.265]	Yes	1.8e-12	2.6e-31	C11orf30-[]--LRRC32
rs10189629	1.146	[1.059,1.241]	Yes	0.00072	1.3e-19	IL1RL2-[]--IL1RL1
rs6906021	1.158	[1.101,1.219]	Yes	1.3e-08	9.6e-23	HLA-DQA1-[]-HLA-DQB1
rs9266772	1.120	[1.043,1.203]	Yes	0.0019	5.5e-15	HLA-C-[]--MICA
rs7720838	1.056	[1.004,1.110]	Yes	0.033	2.7e-12	[]--PTGER4
rs2117339	1.032	[0.983,1.084]	Yes	0.21	1.4e-10	[PLCL1]
rs9860547	1.121	[1.063,1.181]	Yes	2.3e-05	2.7e-14	[LPP]
rs7032572	1.028	[0.962,1.099]	Yes	0.41	7.7e-10	RANBP6-[]--IL33
rs6021270	1.041	[0.940,1.153]	Yes	0.45	3.1e-09	[NFATC2]
rs9303280	1.044	[0.988,1.102]	Yes	0.13	1.0e-09	[GSDMB]
rs962993	1.069	[1.009,1.132]	Yes	0.024	3.3e-10	GATA3-[]
rs17293632	1.039	[0.982,1.099]	Yes	0.18	2.5e-09	[SMAD3]
rs2202749	1.108	[1.051,1.167]	Yes	0.00014	3.4e-12	TPD52-[]--ZBTB10
rs17388568	1.116	[1.058,1.177]	Yes	5.6e-05	2.0e-12	[ADAD1]
rs9671863	1.015	[0.960,1.074]	Yes	0.60	3.1e-08	FOXA1-[]--TTC6
rs13416555	1.121	[1.062,1.183]	Yes	3.3e-05	3.6e-12	[]--ID2
rs7203459	1.033	[0.970,1.101]	Yes	0.31	6.4e-08	[CLEC16A]
rs2107357	1.079	[0.992,1.174]	Yes	0.077	2.4e-08	IL4R-[]--IL21R
rs2056417	1.031	[0.978,1.087]	Yes	0.27	1.4e-07	[PEX14]
rs970924	1.050	[0.993,1.110]	Yes	0.085	5.2e-08	[]--ETS1

OR, odds ratio for the risk allele defined in Table 2; CI, confidence interval; match, agreement of risk alleles across allergy and allergic sensitization studies; P, P value for allergic sensitization; P<sub>fisher</sub>, combined P value across allergy (Table 2) and allergic sensitization meta-analyses using Fisher's method<sup>69</sup>; gene context, as in Table 2.

Supplementary Table 7. NCBI Gap Plus annotations near allergy index SNPs.

allergy association				relationship		Gap Plus annotation					
region	position	SNP	P	distance	r <sup>2</sup>	SNP	P	study ID	MeSH term	genes	
5q22.1	110467499	rs1438673	2.3e-20	-61824	0.771	rs3806932	3.0e-09	20208534	Esophagitis	TSLP	
5q22.1	110467499	rs1438673	2.3e-20	-59497	0.564	rs1898671	5.0e-06	22036096	Rhinitis, Allergic, Seasonal	TSLP	
11q13.5	76299194	rs2155219	1.4e-19	-28511	0.513	rs7130588	2.0e-08	21907864	Asthma	C11orf30 - LRRC32	
11q13.5	76299194	rs2155219	1.4e-19	0	1.000	rs2155219	5.0e-16	21297633	Colitis, Ulcerative	C11orf30 - LRRC32	
11q13.5	76299194	rs2155219	1.4e-19	0	1.000	rs2155219	1.0e-08	22036096	Rhinitis, Allergic, Seasonal	C11orf30 - LRRC32	
11q13.5	76299194	rs2155219	1.4e-19	2122	0.583	rs7927894	1.0e-09	18587394	Crohn Disease	C11orf30 - LRRC32	
11q13.5	76299194	rs2155219	1.4e-19	2122	0.583	rs7927894	8.0e-10	19349984	Dermatitis, Atopic	C11orf30 - LRRC32	
2q12.1	102879464	rs10189629	1.9e-16	74153	0.775	rs3771180	2.0e-15	21804549	Asthma	IL1RL1	
5p13.1	40486896	rs7720838	8.3e-11	3894	0.936	rs10440635	3.0e-07	21743469	Spondylitis, Ankylosing	DAB2 - PTGER4	
2q33.1	198914072	rs10497813	6.1e-10	-17177	0.972	rs6738825	4.0e-09	21102463	Crohn Disease	PLCL1	
3q28	188128979	rs9860547	1.2e-09	-16425	0.698	rs1464510	3.0e-40	20190752	Celiac Disease	LPP	
3q28	188128979	rs9860547	1.2e-09	-16425	0.698	rs1464510	5.0e-09	18311140	Celiac Disease	LPP	
3q28	188128979	rs9860547	1.2e-09	-16425	0.698	rs1464510	1.0e-11	20410501	Vitiligo	LPP	
9p24.1	6172380	rs7032572	1.9e-09	17696	0.949	rs1342326	9.0e-10	20860503	Asthma	RANBP6 - IL33	
9p24.1	6172380	rs7032572	1.9e-09	21075	0.537	rs2381416	2.0e-12	21804549	Asthma	RANBP6 - IL33	
17q12	38074031	rs9303280	8.3e-09	-151772	0.681	rs907092	8.0e-06	19458352	Liver Cirrhosis, Biliary	IKZF3	
17q12	38074031	rs9303280	8.3e-09	-97562	0.781	rs9303277	2.0e-09	20639880	Liver Cirrhosis, Biliary	IKZF3	
17q12	38074031	rs9303280	8.3e-09	-33268	0.745	rs2872507	9.0e-07	20453842	Arthritis, Rheumatoid	ZPBP2 - GSDMB	
17q12	38074031	rs9303280	8.3e-09	-33268	0.745	rs2872507	5.0e-11	21297633	Colitis, Ulcerative	ZPBP2 - GSDMB	
17q12	38074031	rs9303280	8.3e-09	-33268	0.745	rs2872507	5.0e-09	18587394	Crohn Disease	ZPBP2 - GSDMB	
17q12	38074031	rs9303280	8.3e-09	-33268	0.745	rs2872507	2.0e-06	21829393	Diabetes Mellitus, Type 1	ZPBP2 - GSDMB	
17q12	38074031	rs9303280	8.3e-09	-11835	0.752	rs2305480	1.0e-07	20860503	Asthma	GSDMB	
17q12	38074031	rs9303280	8.3e-09	-11835	0.752	rs2305480	3.0e-08	20228799	Colitis, Ulcerative	GSDMB	
17q12	38074031	rs9303280	8.3e-09	-9626	0.760	rs11078927	2.0e-16	21804549	Asthma	GSDMB	
17q12	38074031	rs9303280	8.3e-09	-7791	0.855	rs2290400	6.0e-13	19430480	Diabetes Mellitus, Type 1	GSDMB	
17q12	38074031	rs9303280	8.3e-09	-4082	0.924	rs7216389	9.0e-11	17611496	Asthma	GSDMB	
17q12	38074031	rs9303280	8.3e-09	21143	0.559	rs6503525	5.0e-07	21150878	Asthma	ORMDL3 - GSDMA	
15q22.33	67450305	rs172220858	1.2e-08	-7709	0.989	rs17293632	3.0e-19	21102463	Crohn Disease	SMAD3	
4q27	123329362	rs17388568	3.6e-08	-196870	0.706	rs4505848	5.0e-13	19430480	Diabetes Mellitus, Type 1	KIAA1109	
4q27	123329362	rs17388568	3.6e-08	-130927	0.617	rs6534347	2.0e-06	17554300	Diabetes Mellitus, Type 1	KIAA1109	
4q27	123329362	rs17388568	3.6e-08	0	1.000	rs17388568	9.0e-07	21297633	Colitis, Ulcerative	ADAD1	
4q27	123329362	rs17388568	3.6e-08	0	1.000	rs17388568	6.0e-06	21829393	Diabetes Mellitus, Type 1	ADAD1	
4q27	123329362	rs17388568	3.6e-08	43771	0.909	rs2069772	1.0e-06	22036096	Rhinitis, Allergic, Seasonal	IL2	
8q21.13	81268155	rs6473223	7.7e-08	-449	0.970	rs6992476	9.1e-05	pha002864	Diabetic Nephropathies	RPS5P5 - ZBTB10	
16p13.13	11230703	rs7203459	2.1e-07	-71818	0.508	rs887864	1.0e-06	22036096	Rhinitis, Allergic, Seasonal	CLEC16A	
16p13.13	11230703	rs7203459	2.1e-07	8080	0.596	rs2903692	7.0e-11	17632545	Diabetes Mellitus, Type 1	CLEC16A	

Allergy association, an index SNP from table 2; relationship, the distance and r<sup>2</sup> between the allergy index SNP and the annotation. The Gap Plus annotation includes the reported SNP identifier and P value; a study ID, either a PubMed ID for the reporting study, or a dbGAP “pha” accession number; a MeSH term, or medical subject heading; and associated genes as reported in the original study.

Supplementary Table 8. Nonsynonymous SNPs near allergy index SNPs.

allergy association				relationship		nsSNP annotation		
region	position	SNP	P	distance	r <sup>2</sup>	SNP	gene	mutation
4p14	38811551	rs2101521	5.9e-21	-36662	0.550	rs4129009	TLR10	I775V
4p14	38811551	rs2101521	5.9e-21	-12903	0.676	rs5743618	TLR1	S60I
4p14	38811551	rs2101521	5.9e-21	-11841	0.806	rs4833095	TLR1	N248S
17q12	38074031	rs9303280	8.3e-09	-45397	0.821	rs11557467	ZPB2	S173I
17q12	38074031	rs9303280	8.3e-09	-11835	0.752	rs2305480	GSDMB	P298S
17q12	38074031	rs9303280	8.3e-09	-11814	0.848	rs2305479	GSDMB	G291R

Allergy association, an index SNP from Table 2; relationship, distance and r<sup>2</sup> between the allergy index SNP and the annotation. The nsSNP annotation includes the SNP identifier, affected gene, and amino acid change associated with this SNP.

Supplementary Table 9. Monocyte expression QTLs near allergy index SNPs.

allergy association				relationship		eQTL annotation				
region	position	SNP	P	distance	r <sup>2</sup>	SNP	distance	gene	P	r <sup>2</sup>
4p14	38811551	rs2101521	5.9e-21	-8488	0.711	rs5743592	25160	TLR6	1.4e-08	0.024
2q33.1	198914072	rs10497813	6.1e-10	-33591	0.897	rs1318867	67998	PLCL1	3.0e-28	0.083
3q28	188128979	rs9860547	1.2e-09	-23923	0.872	rs9864529	641581	BCL6	5.2e-10	0.028
17q12	38074031	rs9303280	8.9e-09	-5988	0.924	rs869402	47602	IKZF3	3.0e-07	0.020
2p25.1	8442248	rs10174949	1.0e-07	-1216	0.687	rs6734135	377943	ID2	9.0e-10	0.028
1p36.22	10581658	rs2056417	5.3e-07	-37111	0.861	rs668805	0	PEX14	1.2e-19	0.057

Allergy association, an index SNP from Table 2; relationship, distance and r<sup>2</sup> between the index SNP and the annotation. The eQTL annotation includes the associated SNP; the absolute distance in base pairs between that SNP and the associated gene; the associated gene; the reported P value; and r<sup>2</sup>, the proportion of variance in expression explained by the association.

Supplementary Table 11. Replication of reported associations with asthma

region	pubmed.id	first author	GWAS Catalog result					Allergy result		
			pub date	risk allele	P	OR	genes	P	OR	agree
1q21.3	21804549	Torgerson DG	2011-07-31	rs4845783-G	6.0e-06	CRCT1		0.13	1.02	
1q21.3	21907864	Ferreira MA	2011-09-10	rs4129267-T	2.0e-08	1.09	IL6R	0.034	1.03	+
1q31.3	20032318	Sleiman PM	2009-12-23	rs2786098-C	2.0e-13	1.43	DENND1B, CRB1	0.64	0.99	
2q12.1	21804549	Torgerson DG	2011-07-31	rs3771180-G	2.0e-15		IL1RL1	1.6e-13	1.15	+++
2q12.1	23028483	Ramasamy A	2012-09-28	rs13408661-G	1.0e-09	1.23	IL1RL1, IL18R1	1.9e-15	1.15	+++
2q12.1	20860503	Moffatt MF	2010-09-23	rs3771166-G	3.0e-09	1.15	IL18R1	3.8e-07	1.06	+++
4q31.21	21804548	Hirota T	2011-07-31	rs7686660-T	2.0e-12	1.16	LOC729675	0.099	0.98	
4q31.21	21804548	Hirota T	2011-07-31	rs3805236-G	7.0e-08	1.12	GAB1	0.00031	1.06	+++
5q12.1	19426955	Himes BE	2009-05-07	rs1588265-C	3.0e-08	1.18	PDE4D	0.26	1.01	
5q22.1	21804549	Torgerson DG	2011-07-31	rs1837253-C	1.0e-14		TSLP	2.3e-10	1.09	+++
5q22.1	21804548	Hirota T	2011-07-31	rs1837253-C	1.0e-16	1.17	TSLP	2.3e-10	1.09	+++
5q31.1	20860503	Moffatt MF	2010-09-23	rs2073643-T	2.0e-07	1.11	SLC22A5	0.090	1.02	
5q31.1	20159242	Li X	2010-02-01	rs2244012-C	3.0e-07	1.64	RAD50	0.00056	1.05	++
5q31.1	20860503	Moffatt MF	2010-09-23	rs1295686-T	1.0e-07	1.15	IL13	1.7e-06	1.07	+++
6p21.32	21804548	Hirota T	2011-07-31	rs204993-A	2.0e-15	1.17	PBX2	7.3e-07	0.94	---
6p21.32	21804548	Hirota T	2011-07-31	rs404860-A	4.0e-23	1.21	NOTCH4	0.61	1.01	
6p21.32	21804548	Hirota T	2011-07-31	rs3129943-T	3.0e-15	1.17	C6orf10	0.20	1.02	
6p21.32	21804548	Hirota T	2011-07-31	rs3117098-G	5.0e-12	1.16	BTNL2	0.45	1.01	
6p21.32	23028483	Ramasamy A	2012-09-28	rs9268516-T	1.0e-08	1.15	BTNL2, HLA-DRA	4.2e-05	1.06	+++
6p21.32	21804548	Hirota T	2011-07-31	rs3129890-T	5.0e-13	1.15	HLA-DRA	0.56	1.01	
6p21.32	21804548	Hirota T	2011-07-31	rs7775228-A	5.0e-15	1.17	HLA-DQB1	0.00012	0.94	---
6p21.32	21804548	Hirota T	2011-07-31	rs9275698-T	5.0e-12	1.18	HLA-DQA2	0.68	1.01	
6p21.32	21804548	Hirota T	2011-07-31	rs9500927-T	4.0e-09	1.13	HLA-DOA	0.0062	1.05	+
6p21.32	21814517	Noguchi E	2011-07-21	rs987870-C	2.0e-10	1.40	HLA, DPB1	9.2e-05	1.07	+++
8q24.11	21814517	Noguchi E	2011-07-21	rs3019885-G	5.0e-13	1.34	SLC30A8	0.068	0.98	
9p24.1	20860503	Moffatt MF	2010-09-23	rs1342326-C	9.0e-10	1.20	IL33	6.0e-09	1.10	+++
9p24.1	21804549	Torgerson DG	2011-07-31	rs2381416-C	2.0e-12		IL33	1.2e-07	1.07	+++
10p14	21804548	Hirota T	2011-07-31	rs10508372-C	2.0e-15	1.16	LOC338591	0.00019	1.09	+++
10q21.1	21907864	Ferreira MA	2011-09-10	rs7922491-A	5.0e-07	1.13	PRKG1	0.46	1.01	
11q13.5	21907864	Ferreira MA	2011-09-10	rs7130588-G	2.0e-08	1.09	LRRK32	2.2e-12	1.09	+++
12q13.2	21804548	Hirota T	2011-07-31	rs2069408-C	1.0e-10	1.15	CDK2	0.060	1.02	
12q13.2	21804548	Hirota T	2011-07-31	rs1701704-G	2.0e-13	1.19	IKZF4	0.012	1.03	+
15q22.2	20860503	Moffatt MF	2010-09-23	rs11071559-C	1.0e-07	1.18	RORA	9.2e-05	1.07	+++
15q22.33	20860503	Moffatt MF	2010-09-23	rs744910-G	4.0e-09	1.12	SMAD3	4.2e-06	1.06	+++
17q12	20860503	Moffatt MF	2010-09-23	rs2305480-G	1.0e-07	1.18	GSDMB	4.4e-08	1.07	+++
17q12	21804549	Torgerson DG	2011-07-31	rs11078927-C	2.0e-16		GSDMB	3.9e-08	0.94	+++
17q12	17611496	Moffatt MF	2007-07-26	rs7216389-T	9.0e-11	1.45	ORMDL3	3.2e-08	1.07	+++
17q12	21150878	Ferreira MA	2010-12-08	rs6503525-C	5.0e-07	1.33	ORMDL3	5.9e-06	1.06	+++
17q21.1	20860503	Moffatt MF	2010-09-23	rs3894194-A	5.0e-09	1.17	GSDMA	7.9e-06	1.05	+++
22q12.3	20860503	Moffatt MF	2010-09-23	rs2284033-G	1.0e-08	1.12	IL2RB	0.016	1.03	+

Results are shown for all asthma reports in the NHGRI GWAS Catalog for which the risk allele was available or could be inferred from the original study, and the reported SNP was tested in our allergy analysis. Entries in gray italics overlap a region identified in Table 2. Allergy result: *P*, *P* value from shared-effect meta-analysis; OR, odds ratio for the previously reported risk allele; agree, '+' *P*<0.05, '++' *P*<0.005, '+++' *P*<0.0005 with the same risk allele; dashes indicate the same levels of significance but with inconsistent risk alleles.

Supplementary Table 12. Replication of reported associations with atopic dermatitis.

region	pubmed.id	first author	GWAS Catalog result					Allergy result		
			pubdate	riskallele	P	OR	genes	P	OR	agree
1q21.3	21666691	Sun LD	2011-06-13	rs3126085-A	6.0e-12	1.22	FLG	0.068	0.97	
<i>2q12.1</i>	<i>23042114</i>	<i>Hirota T</i>	<i>2012-10-07</i>	<i>rs13015714-G</i>	<i>8.0e-18</i>	<i>1.27</i>	<i>IL1RL1, IL18R1, IL18RAP</i>	<i>0.18</i>	<i>0.98</i>	
2q13	23042114	Hirota T	2012-10-07	rs2271404-G	3.0e-07	1.17	LOC100505634	0.58	0.99	
3p22.3	23042114	Hirota T	2012-10-07	rs7613051-A	6.0e-21	1.29	GLB1	0.19	1.02	
3q13.2	23042114	Hirota T	2012-10-07	rs12634229-G	2.0e-19	1.29	CCDC80, LOC100630917	0.49	1.02	
5q31.1	23042114	Hirota T	2012-10-07	rs1295686-T	2.0e-06	1.22	KIF3A, IL13	1.7e-06	1.07	+++
5q31.1	22197932	Paternoster L	2011-12-25	rs2897442-C	4.0e-08	1.11	KIF3A	0.024	1.03	+
<i>6p21.33</i>	<i>23042114</i>	<i>Hirota T</i>	<i>2012-10-07</i>	<i>rs9368677-G</i>	<i>1.0e-17</i>	<i>1.36</i>	<i>HLA-C</i>	<i>0.12</i>	<i>0.96</i>	
<i>6p21.33</i>	<i>22197932</i>	<i>Paternoster L</i>	<i>2011-12-25</i>	<i>rs3853601-G</i>	<i>2.0e-06</i>	<i>1.13</i>	<i>BAT1</i>	<i>0.92</i>	<i>1.00</i>	
<i>6p21.32</i>	<i>23042114</i>	<i>Hirota T</i>	<i>2012-10-07</i>	<i>rs176095-T</i>	<i>8.0e-20</i>	<i>1.40</i>	<i>GPSM3</i>	<i>9.6e-07</i>	<i>0.93</i>	---
<i>6p21.32</i>	<i>23042114</i>	<i>Hirota T</i>	<i>2012-10-07</i>	<i>rs9469099-G</i>	<i>5.0e-19</i>	<i>1.61</i>	<i>C6orf10</i>	<i>0.35</i>	<i>1.04</i>	
7p22.2	23042114	Hirota T	2012-10-07	rs4722404-G	8.0e-09	1.18	CARD11	0.0069	1.03	+
8q24.21	23042114	Hirota T	2012-10-07	rs7815944-A	4.0e-07	1.16	MIR1208	0.0033	1.10	++
10q21.2	23042114	Hirota T	2012-10-07	rs10995251-C	6.0e-20	1.28	ZNF365	0.090	1.02	
11p15.4	23042114	Hirota T	2012-10-07	rs878860-G	2.0e-22	1.31	OR10A3, NLRP10	0.082	0.98	
11q13.1	23042114	Hirota T	2012-10-07	rs593982-C	6.0e-07	1.23	OVOL1	0.47	0.99	
11q13.1	22197932	Paternoster L	2011-12-25	rs479844-G	1.0e-13	1.14	OVOL1	0.15	1.02	
<i>11q13.5</i>	<i>19349984</i>	<i>Esparza-Gordillo</i>	<i>2009-04-06</i>	<i>rs7927894-A</i>	<i>8.0e-10</i>	<i>1.22</i>	<i>C11orf30</i>	<i>6.2e-16</i>	<i>1.10</i>	+++
<i>11q13.5</i>	<i>23042114</i>	<i>Hirota T</i>	<i>2012-10-07</i>	<i>rs11236809-G</i>	<i>3.0e-06</i>	<i>1.24</i>	<i>C11orf30</i>	<i>0.20</i>	<i>1.03</i>	
<i>16p13.13</i>	<i>23042114</i>	<i>Hirota T</i>	<i>2012-10-07</i>	<i>rs9923856-A</i>	<i>6.0e-06</i>	<i>1.17</i>	<i>CLEC16A</i>	<i>4.4e-05</i>	<i>1.05</i>	+++
19p13.2	22197932	Paternoster L	2011-12-25	rs2164983-A	7.0e-09	1.16	ACTL9	0.20	1.02	
20q13.2	23042114	Hirota T	2012-10-07	rs16999165-T	2.0e-08	1.19	CYP24A1, PFDN4	0.24	1.04	
22q12.3	22197932	Paternoster L	2011-12-25	rs4821544-C	6.0e-06	1.09	NCF4	0.0064	1.04	+

Results are shown for all atopic dermatitis reports in the NHGRI GWAS Catalog for which the risk allele was available or could be inferred from the original study, and the reported SNP was tested in our allergy analysis. Entries in gray italics overlap a region identified in Table 2. Allergy result: *P*, *P* value from shared-effect meta-analysis; OR, odds ratio for the previously reported risk allele; agree, '+'  $P<0.05$ , '++'  $P<0.005$ , '+++'  $P<0.0005$  with the same risk allele; dashes indicate the same levels of significance but with inconsistent risk alleles.

Supplementary Table 13. Association of index SNPs with symptoms in the 23andMe cohort.

SNP	OR	95% CI	$P_{main}$	$P_{int}$	$P_{total}$	gene context
rs2101521	1.143	[1.109,1.178]	2.6e-18	0.62	1.2e-16	TLR1-[]--TLR6
rs1438673	1.124	[1.097,1.152]	1.8e-20	0.34	5.7e-19	WDR36-[]--CAMK4
rs2155219	1.106	[1.079,1.133]	1.1e-15	0.17	1.3e-14	C11orf30-[]--LRRC32
rs10189629	1.170	[1.128,1.213]	1.6e-17	0.20	2.6e-16	IL1RL2-[]--IL1RL1
rs6906021	1.105	[1.078,1.133]	4.5e-15	0.064	1.8e-14	HLA-DQA1-[]-HLA-DQB1
rs9266772	1.109	[1.075,1.143]	6.4e-11	0.91	2.6e-09	HLA-C---[]--MICA
rs7720838	1.076	[1.050,1.103]	7.9e-09	0.96	2.7e-07	[]--PTGER4
rs10497813	1.067	[1.041,1.093]	2.2e-07	0.0078	5.5e-08	[PLCL1]
rs9860547	1.080	[1.053,1.107]	2.0e-09	0.54	4.3e-08	[LPP]
rs7032572	1.142	[1.097,1.188]	6.4e-11	0.043	1.3e-10	RANBP6---[]--IL33
rs6021270	1.167	[1.105,1.234]	3.6e-08	0.69	8.5e-07	[NFATC2]
rs9303280	1.064	[1.038,1.090]	1.0e-06	3.5e-05	1.2e-09	[GSDMB]
rs17228058	1.096	[1.065,1.128]	3.2e-10	0.13	1.8e-09	[SMAD3]
rs962993	1.068	[1.041,1.096]	4.5e-07	0.32	3.8e-06	GATA3---[]
rs17388568	1.077	[1.048,1.107]	8.2e-08	0.30	7.2e-07	[ADAD1]
rs1998359	1.079	[1.046,1.112]	1.2e-06	0.024	7.8e-07	FOXA1-[]--TTC6
rs6473223	1.070	[1.042,1.097]	2.7e-07	0.55	4.2e-06	TPD52---[]--ZBTB10
rs10174949	1.058	[1.030,1.088]	5.1e-05	0.65	0.00063	[]--ID2
rs7203459	1.079	[1.049,1.110]	1.0e-07	0.081	2.7e-07	[CLEC16A]
rs2107357	1.078	[1.040,1.116]	3.2e-05	0.76	0.00046	IL4R-[]--IL21R
rs2056417	1.064	[1.036,1.092]	5.5e-06	0.22	2.9e-05	[PEX14]
rs10893845	1.064	[1.038,1.090]	7.2e-07	0.20	3.7e-06	[]--ETS1
rs17533090	1.122	[1.085,1.161]	2.7e-11	0.32	4.1e-10	

OR, odds ratio per copy of the risk allele, for the main effect across allergy types; CI, confidence interval;  $P_{main}$ , test of main effect across symptoms;  $P_{int}$ , test for interaction with symptom;  $P_{total}$ , joint test including main effect and interaction. The across-symptom test

Supplementary Table 14. SNP effects on individual symptoms in the 23andMe cohort.

SNP	Rhinitis			Asthma			Contact Dermatitis		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
rs2101521	1.149	[1.112,1.188]	6.0e-17	1.140	[1.089,1.192]	1.1e-08	1.125	[1.076,1.177]	1.8e-07
rs1438673	1.131	[1.101,1.162]	3.7e-19	1.126	[1.084,1.169]	3.1e-10	1.102	[1.063,1.143]	1.2e-07
rs2155219	1.103	[1.073,1.133]	9.7e-13	1.133	[1.092,1.175]	2.1e-11	1.091	[1.052,1.131]	1.9e-06
rs10189629	1.167	[1.122,1.214]	9.9e-15	1.210	[1.144,1.279]	1.0e-11	1.145	[1.084,1.209]	9.2e-07
rs6906021	1.110	[1.080,1.141]	7.1e-14	1.124	[1.083,1.167]	4.3e-10	1.073	[1.034,1.113]	0.00015
rs9266772	1.111	[1.074,1.150]	1.2e-09	1.109	[1.059,1.162]	8.6e-06	1.101	[1.052,1.152]	2.8e-05
rs7720838	1.078	[1.048,1.108]	6.9e-08	1.074	[1.035,1.115]	0.00015	1.073	[1.034,1.113]	0.00016
rs10497813	1.081	[1.052,1.110]	1.6e-08	1.072	[1.033,1.112]	0.00017	1.023	[0.986,1.061]	0.21
rs9860547	1.080	[1.051,1.111]	3.5e-08	1.065	[1.026,1.106]	0.00077	1.090	[1.051,1.132]	3.8e-06
rs7032572	1.140	[1.091,1.192]	4.1e-09	1.196	[1.127,1.268]	1.5e-09	1.100	[1.037,1.166]	0.0014
rs6021270	1.167	[1.100,1.239]	2.8e-07	1.193	[1.096,1.298]	3.5e-05	1.146	[1.056,1.245]	0.0010
rs9303280	1.046	[1.018,1.075]	0.0011	1.133	[1.091,1.176]	3.3e-11	1.056	[1.018,1.096]	0.0032
rs17228058	1.096	[1.062,1.132]	1.1e-08	1.125	[1.078,1.174]	3.9e-08	1.072	[1.027,1.119]	0.0012
rs962993	1.067	[1.038,1.098]	4.9e-06	1.088	[1.047,1.130]	1.3e-05	1.053	[1.014,1.094]	0.0071
rs17388568	1.069	[1.037,1.102]	1.2e-05	1.101	[1.058,1.146]	1.8e-06	1.079	[1.036,1.123]	0.00018
rs1998359	1.074	[1.038,1.111]	3.6e-05	1.126	[1.076,1.178]	2.2e-07	1.053	[1.006,1.102]	0.025
rs6473223	1.065	[1.036,1.096]	9.8e-06	1.086	[1.046,1.129]	1.7e-05	1.067	[1.027,1.108]	0.00070
rs10174949	1.056	[1.024,1.088]	0.00041	1.074	[1.030,1.119]	0.00070	1.054	[1.013,1.098]	0.0095
rs7203459	1.073	[1.041,1.107]	5.4e-06	1.117	[1.070,1.165]	2.5e-07	1.064	[1.020,1.109]	0.0032
rs2107357	1.071	[1.030,1.114]	0.00050	1.086	[1.030,1.144]	0.0018	1.089	[1.035,1.147]	0.00097
rs2056417	1.070	[1.039,1.102]	4.9e-06	1.037	[0.996,1.079]	0.072	1.068	[1.027,1.111]	0.00089
rs10893845	1.072	[1.043,1.101]	3.7e-07	1.066	[1.028,1.106]	0.00051	1.039	[1.001,1.077]	0.039
rs17533090	1.131	[1.089,1.174]	7.3e-11	1.129	[1.072,1.189]	2.9e-06	1.091	[1.037,1.147]	0.00064

OR, odds ratio per copy of the risk allele, for specific self reported allergy symptom; CI, confidence interval; P, Wald test for significance of the effect for this particular symptom.

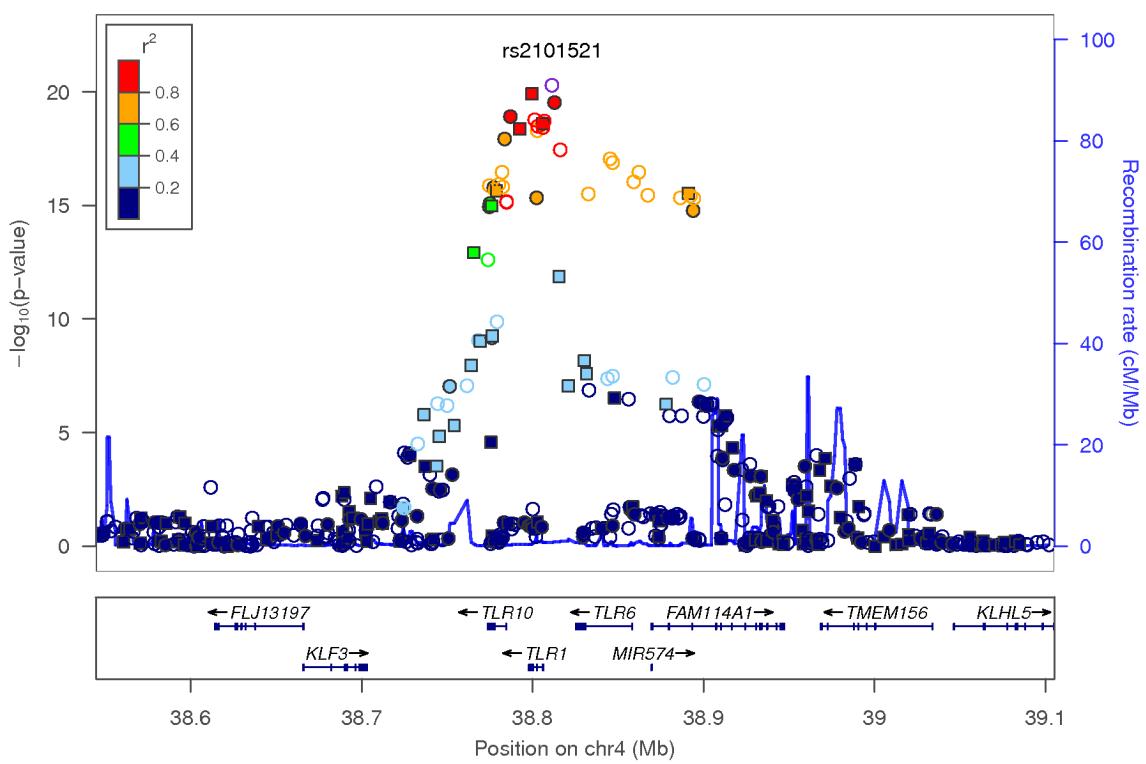
Supplementary Table 15. Consistency of effects at shared allergy and autoimmune loci

region	gene	diseases	match
11q13.5	<i>LRRK32</i>	CD, UC	yes
5p13.1	<i>PTGER4</i>	AS	yes
2q33.1	<i>PLCL1</i>	CD	yes
3q28	<i>LPP</i>	CE, VI	no
17q12	<i>GSDMB</i>	BC, CD, UC, RA, T1D	no
15q22.33	<i>SMAD3</i>	CD	yes
4q27	<i>ADAD1</i>	UC, T1D	yes
16p13.13	<i>CLEC16A</i>	T1D	yes

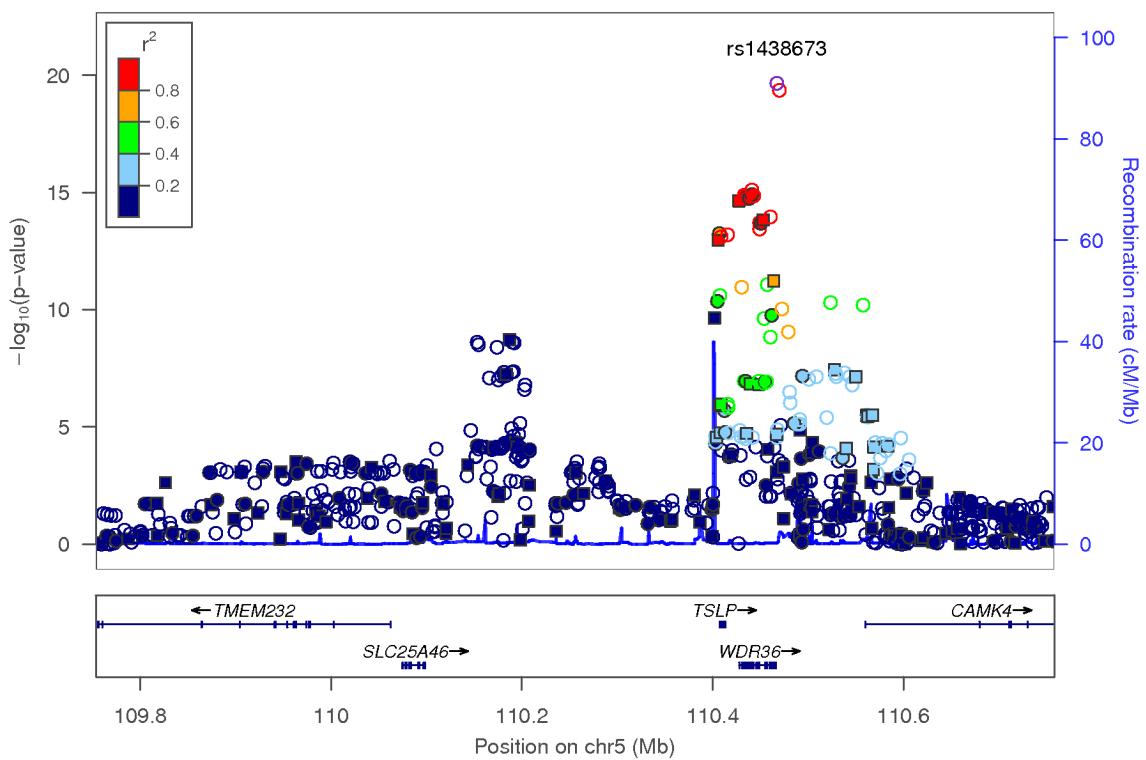
Diseases: CD, Crohn's disease; UC, ulcerative colitis; BC, primary biliary cirrhosis; RA, rheumatoid arthritis; T1D, type 1 diabetes; CE, celiac disease; VI, vitiligo; AS, ankylosing spondylitis. Match, indicates whether allergy risk allele matches risk allele(s) for autoimmune disease.

Supplementary Figure 1. Regional plots of allergy associations. Symbol colors indicate linkage disequilibrium with the index SNP, which is labeled and colored purple. Open circles indicate imputed variants, filled circles indicate partially genotyped variants, and filled squares indicate fully genotyped variants. Plots were generated using LocusZoom<sup>73</sup>. Results are in NCBI Build 37 coordinates.

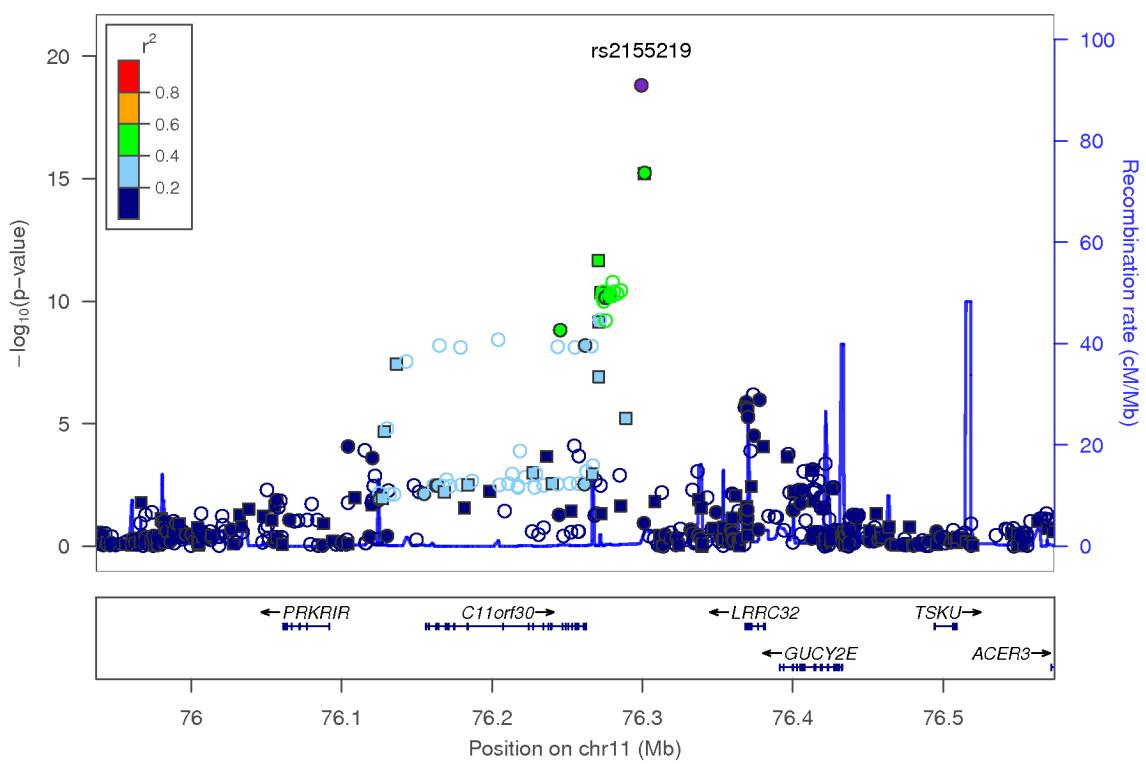
A



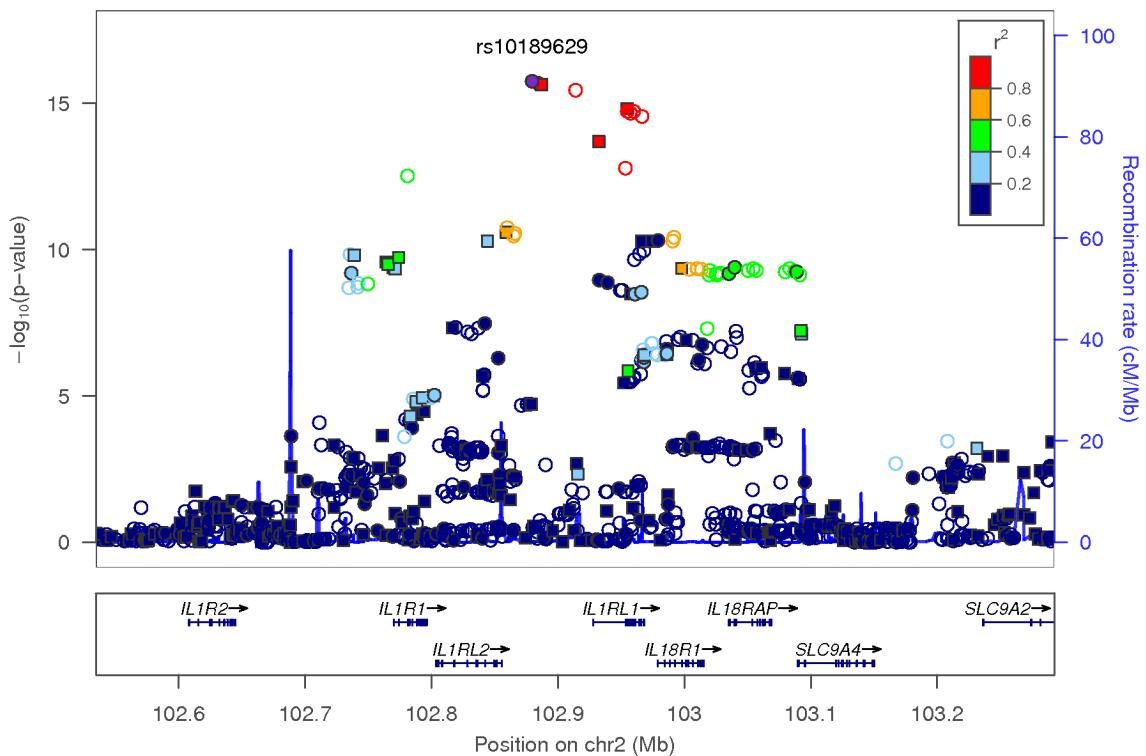
B



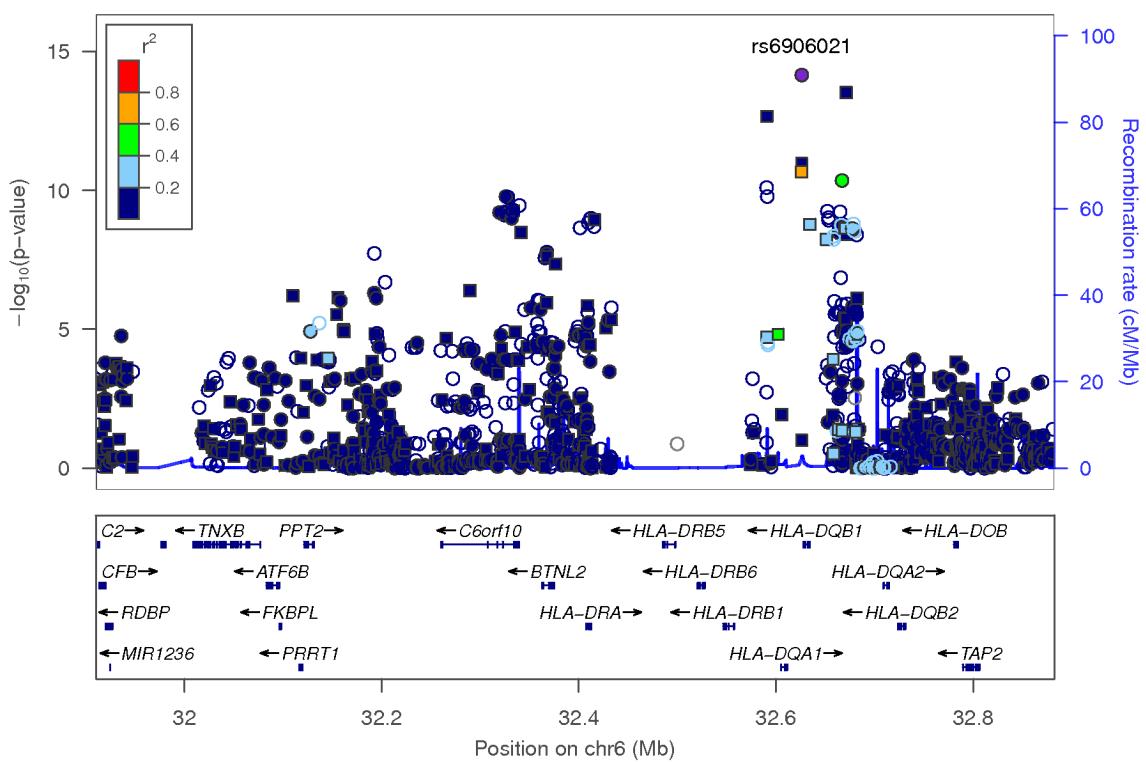
C



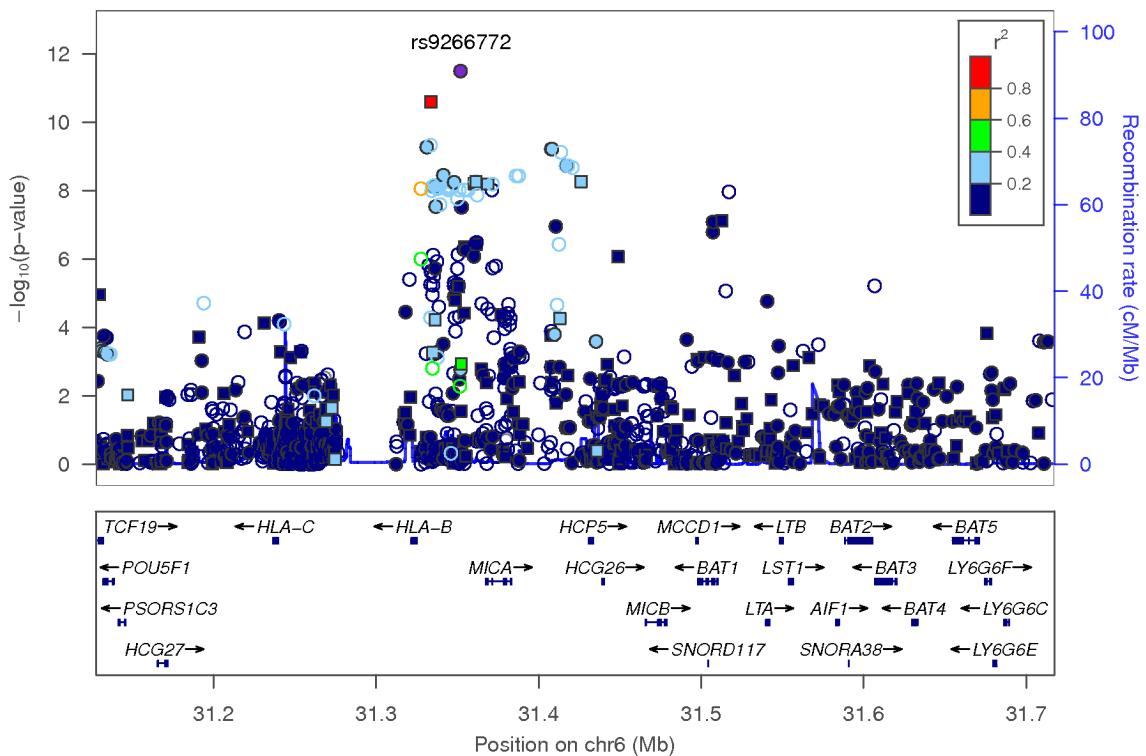
D



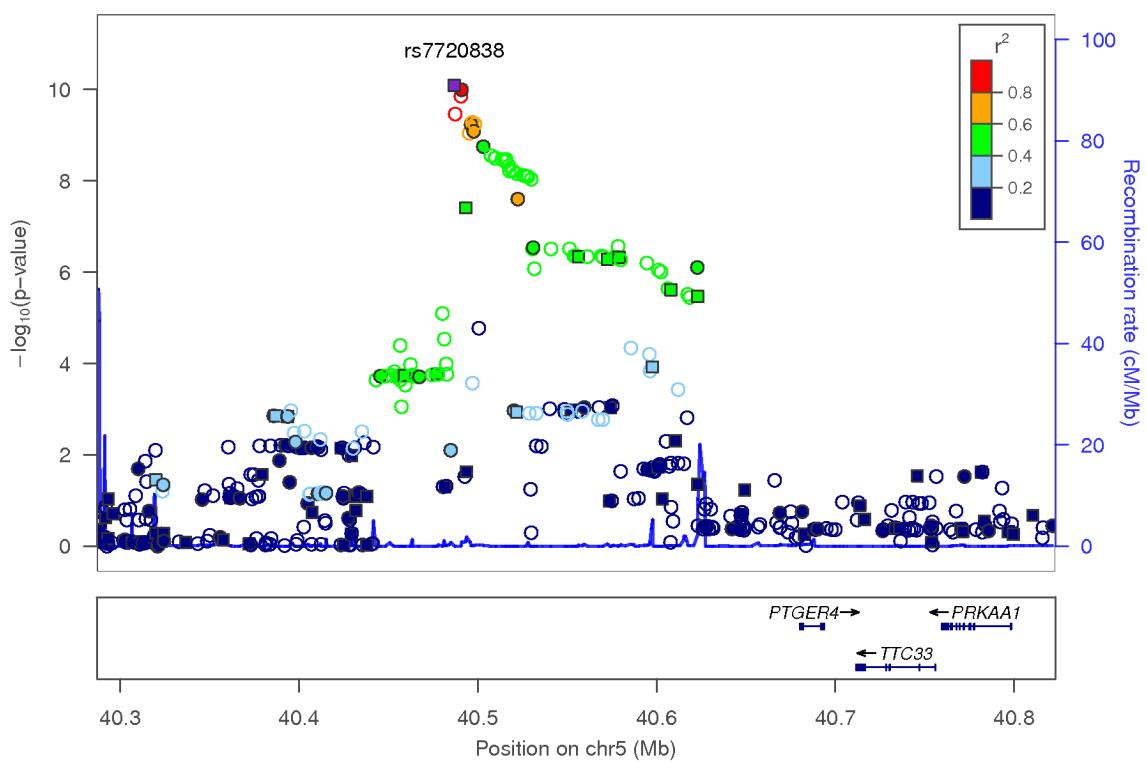
E



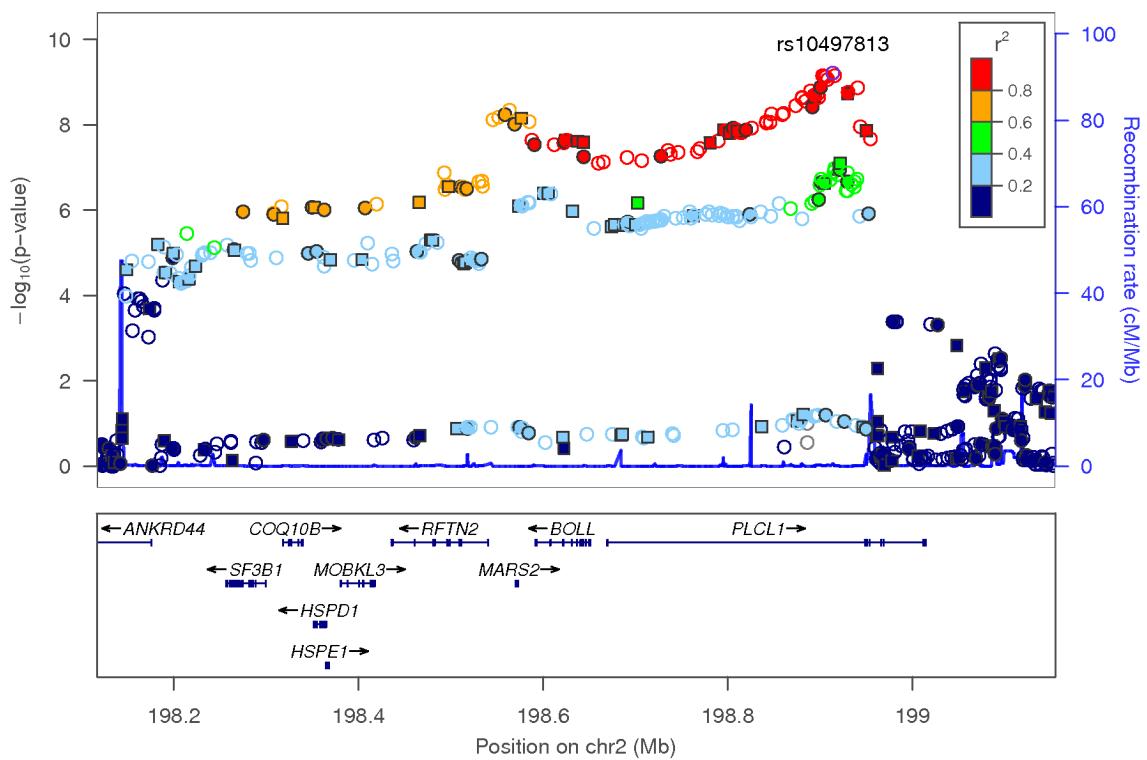
F



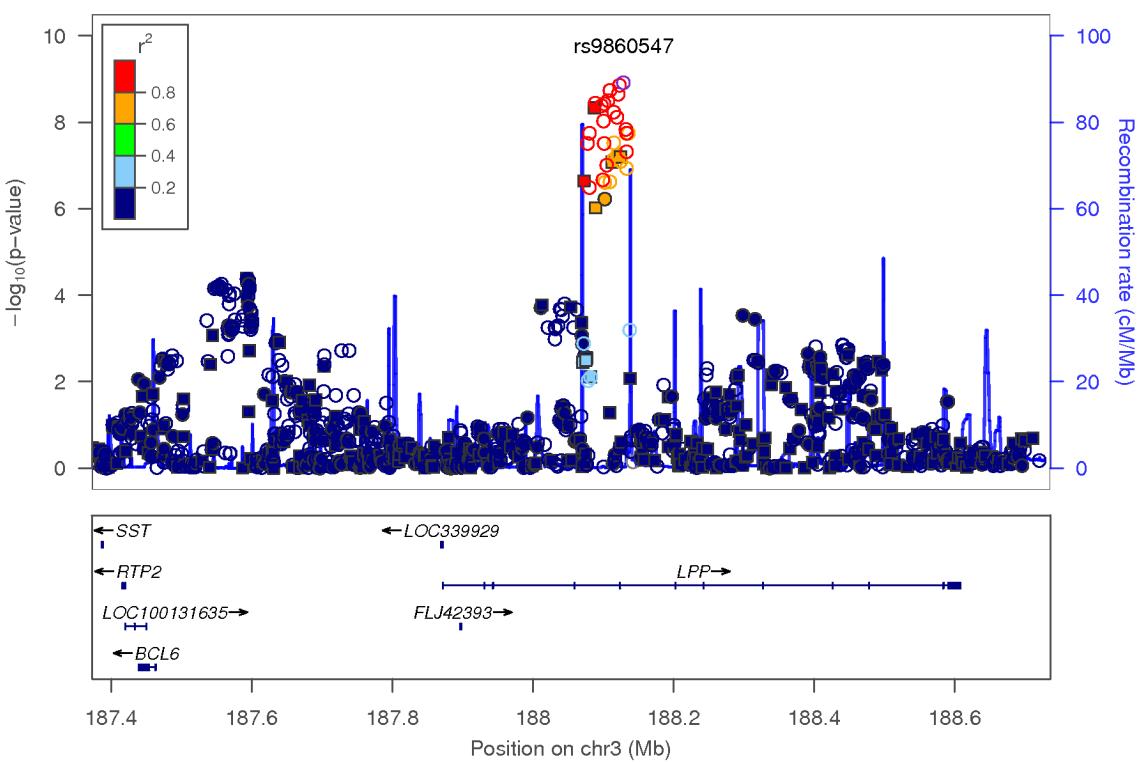
G



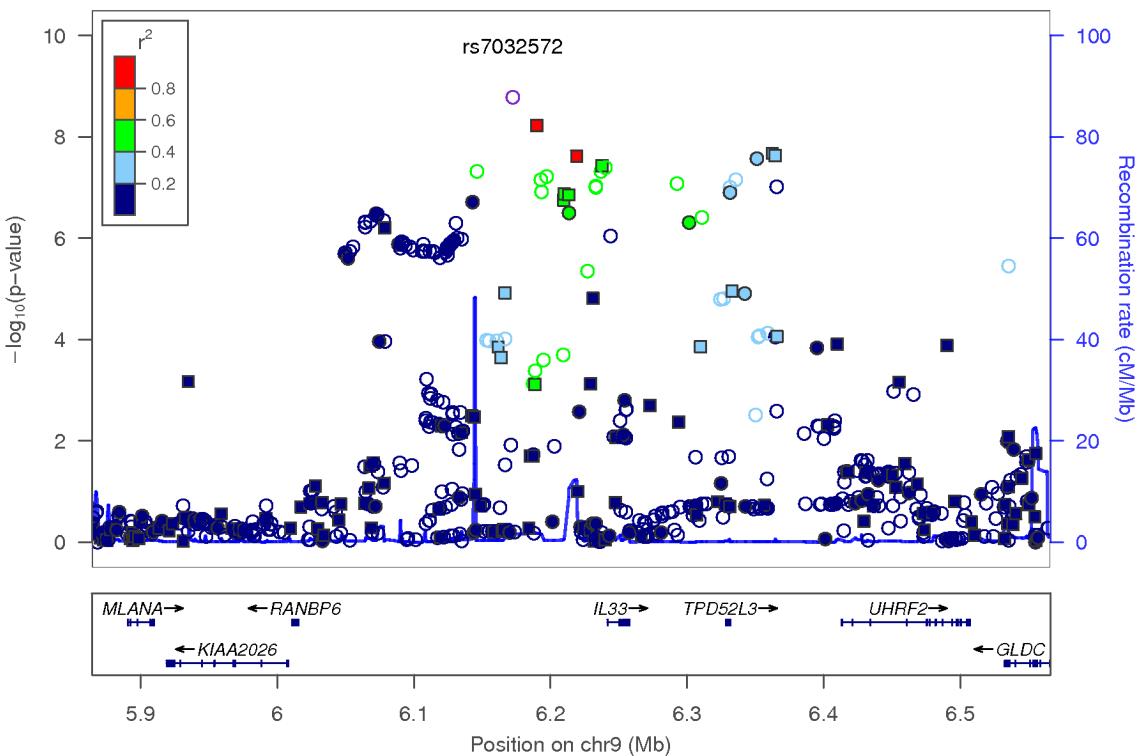
H



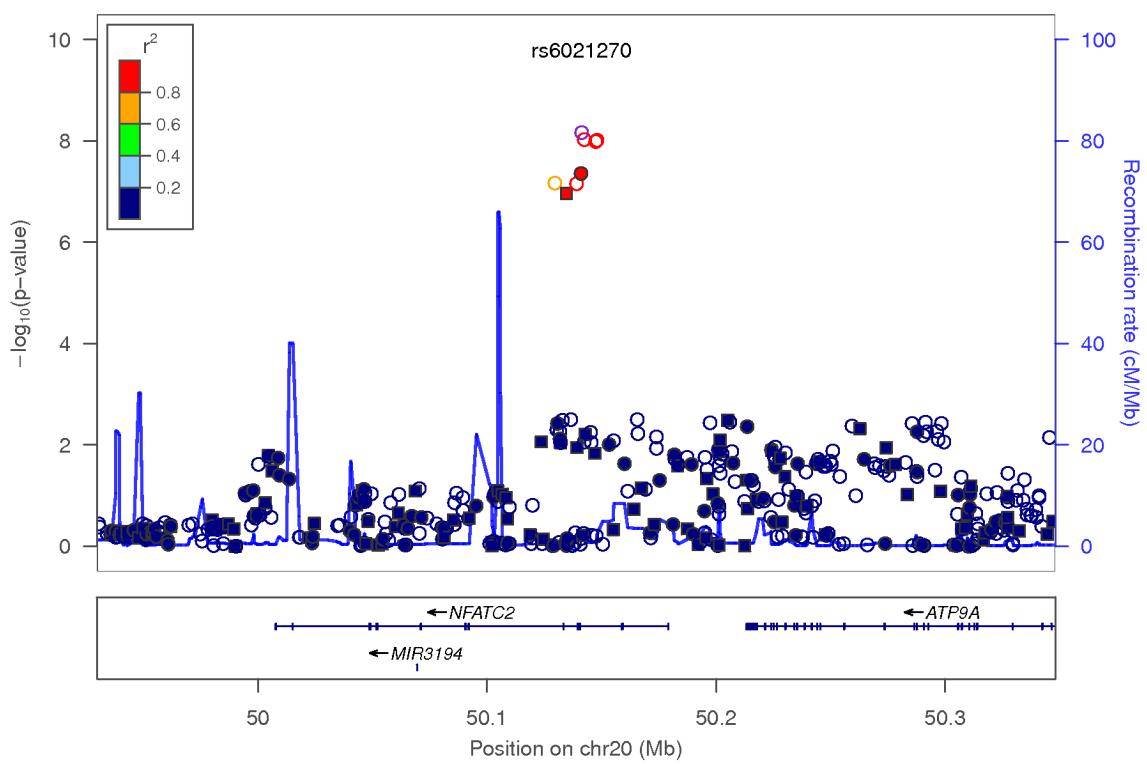
I



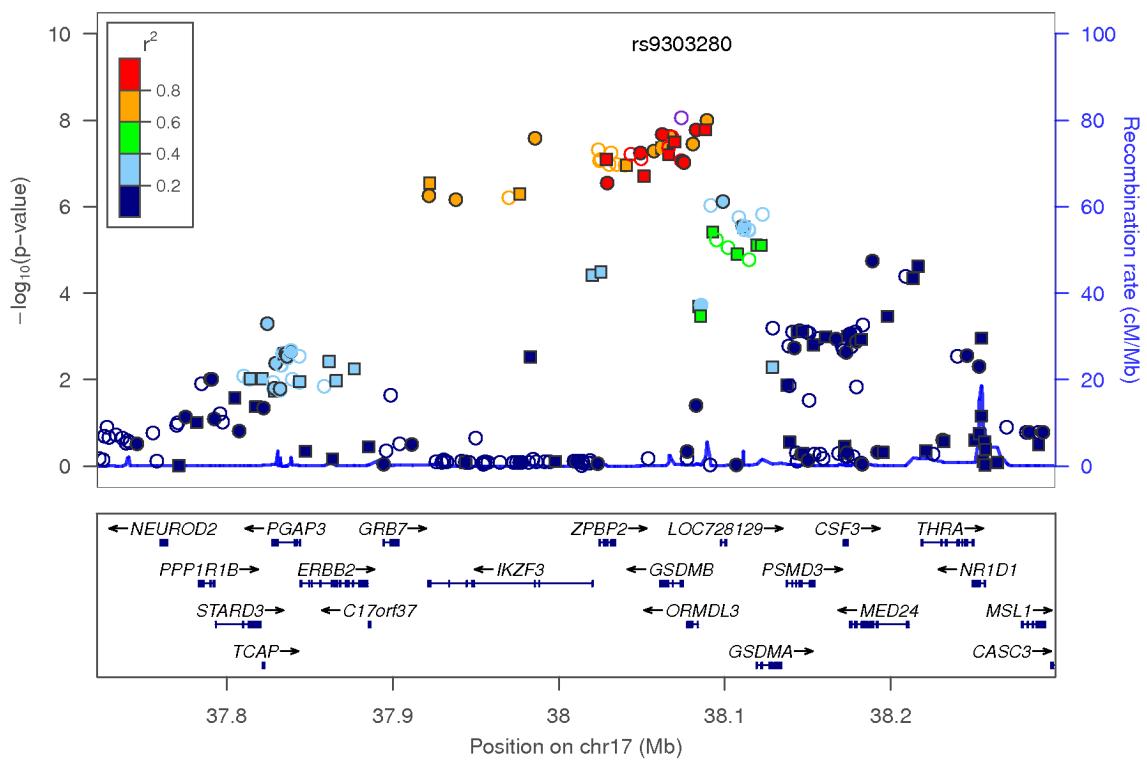
J



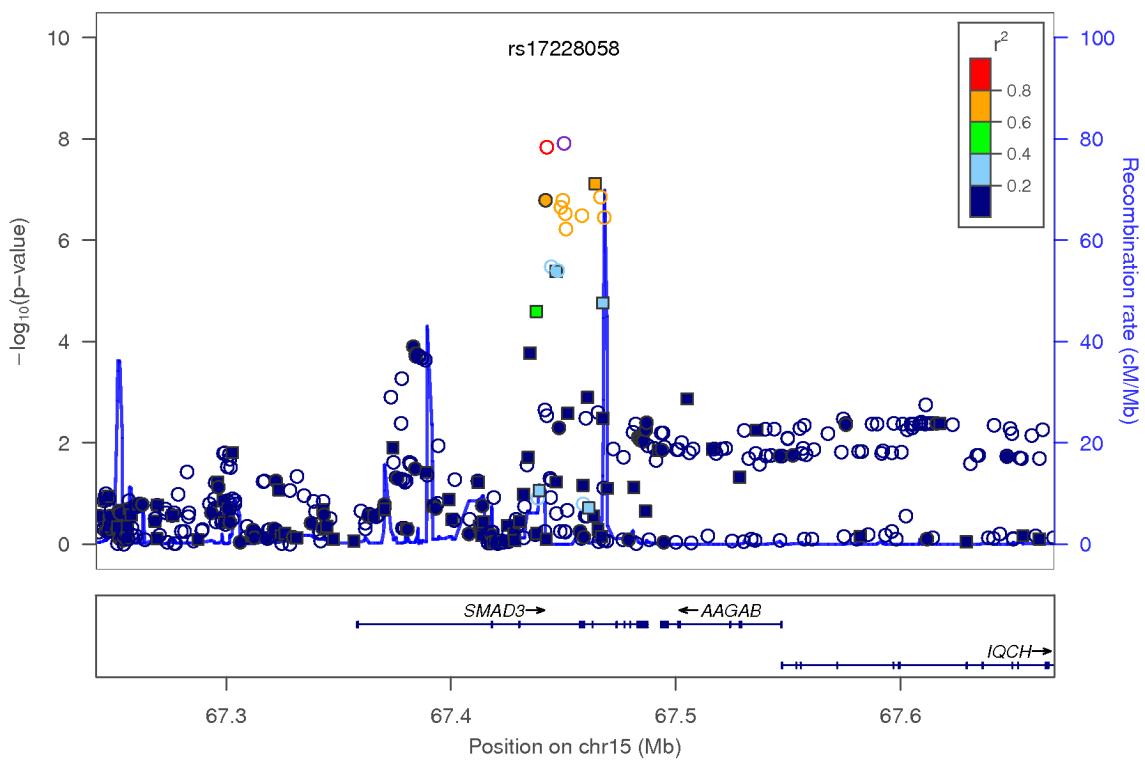
K



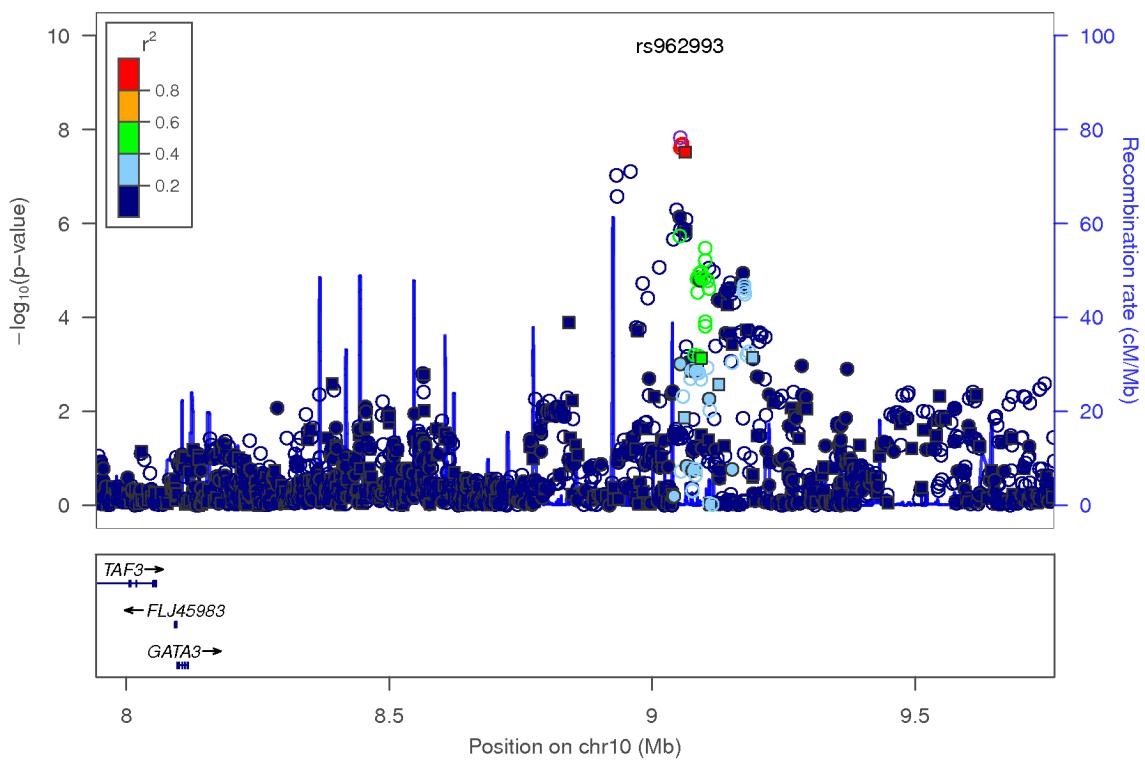
L



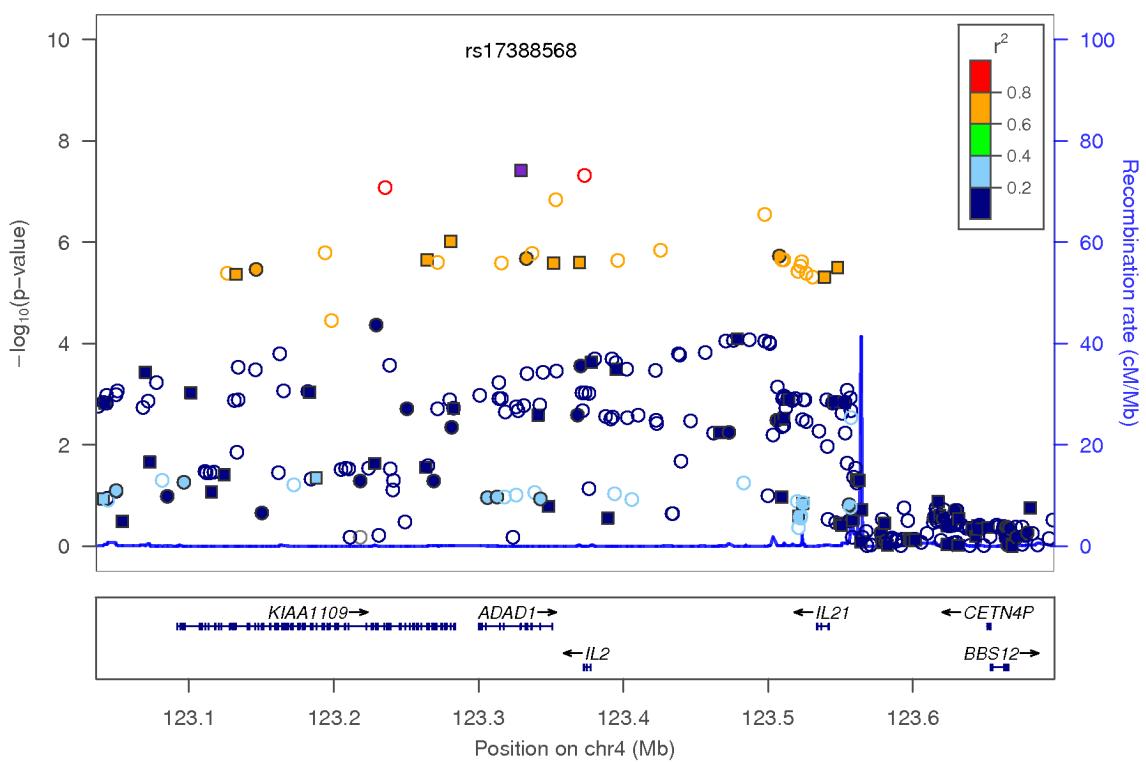
M



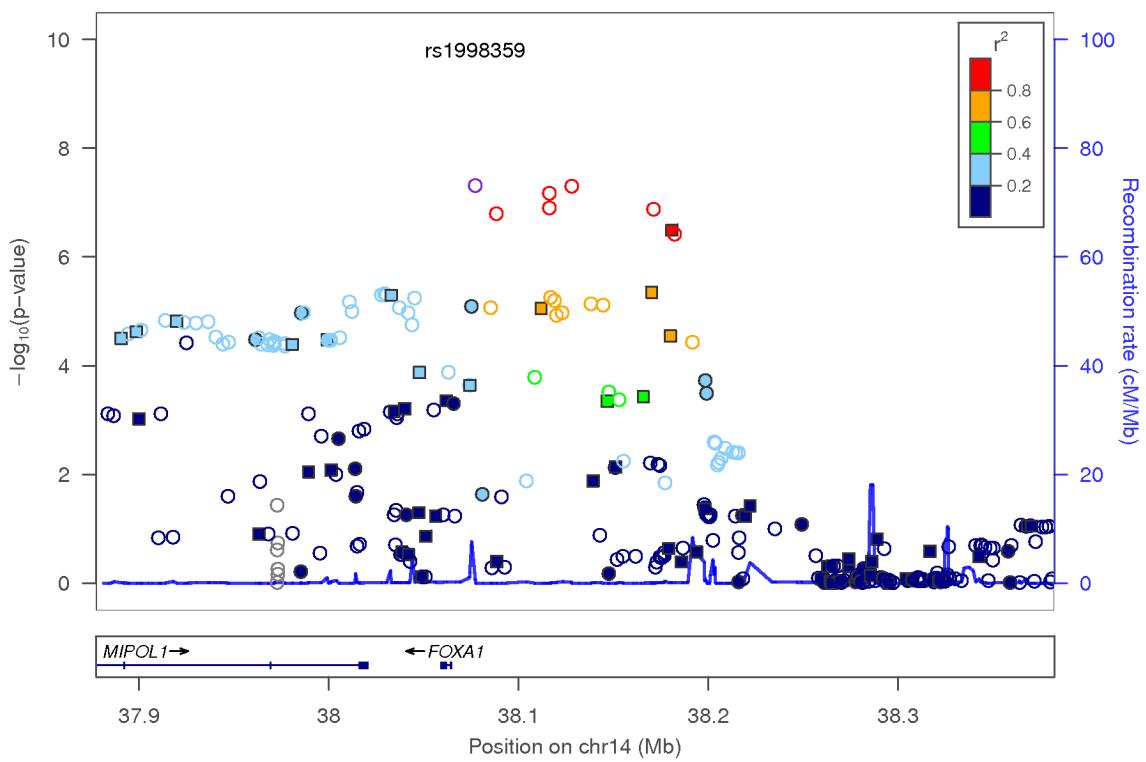
N



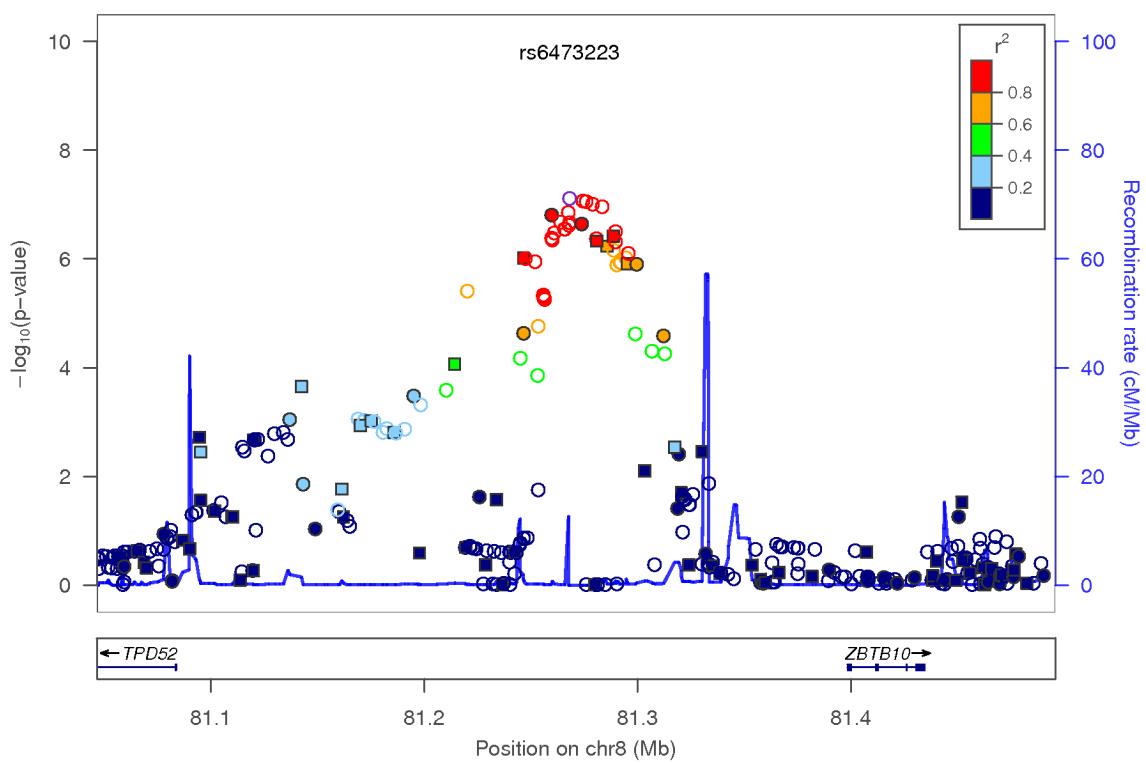
0



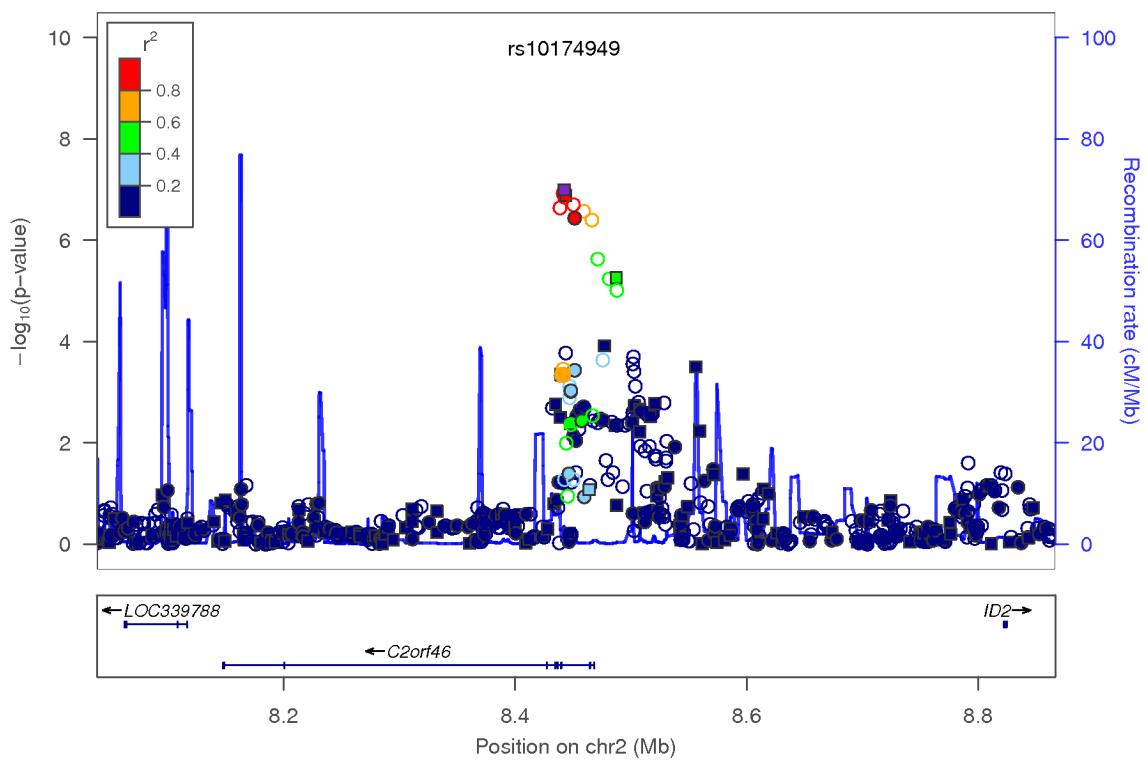
P



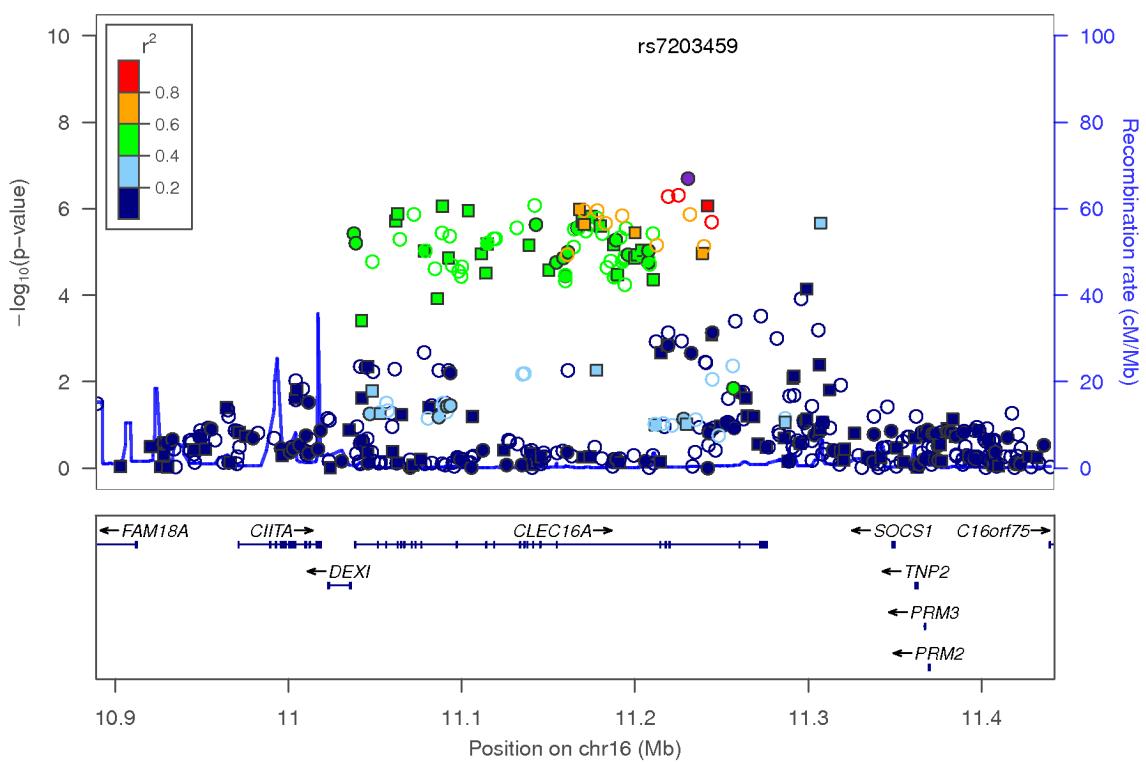
Q



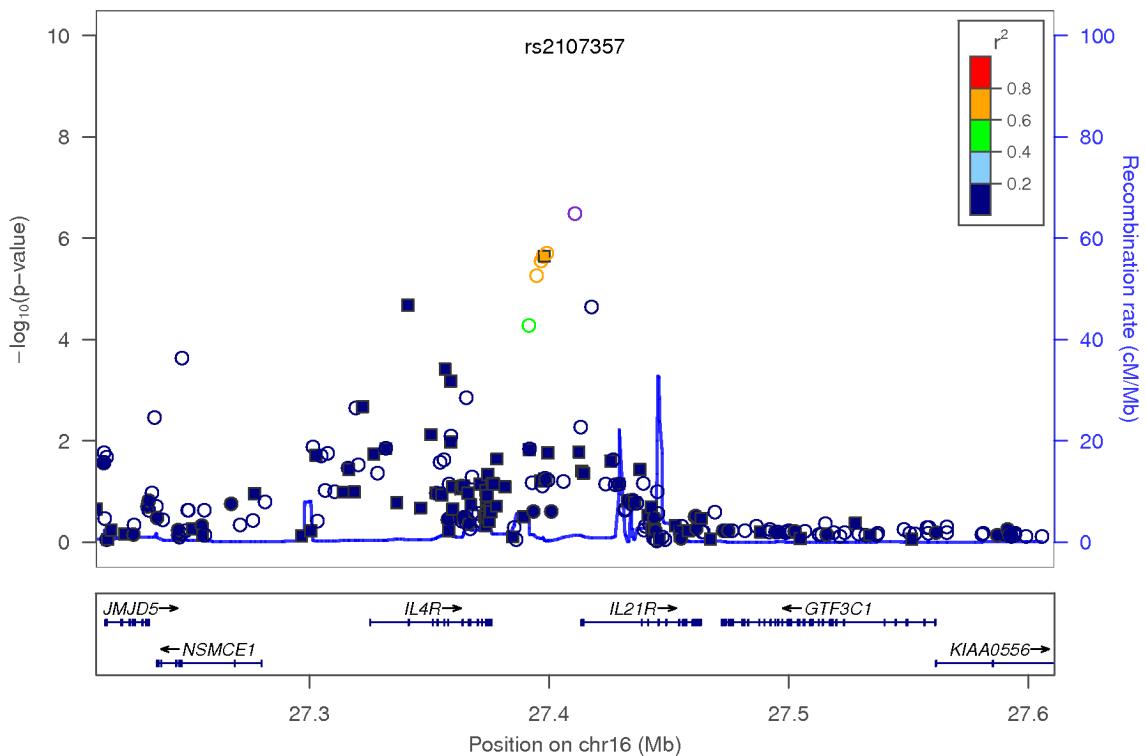
R



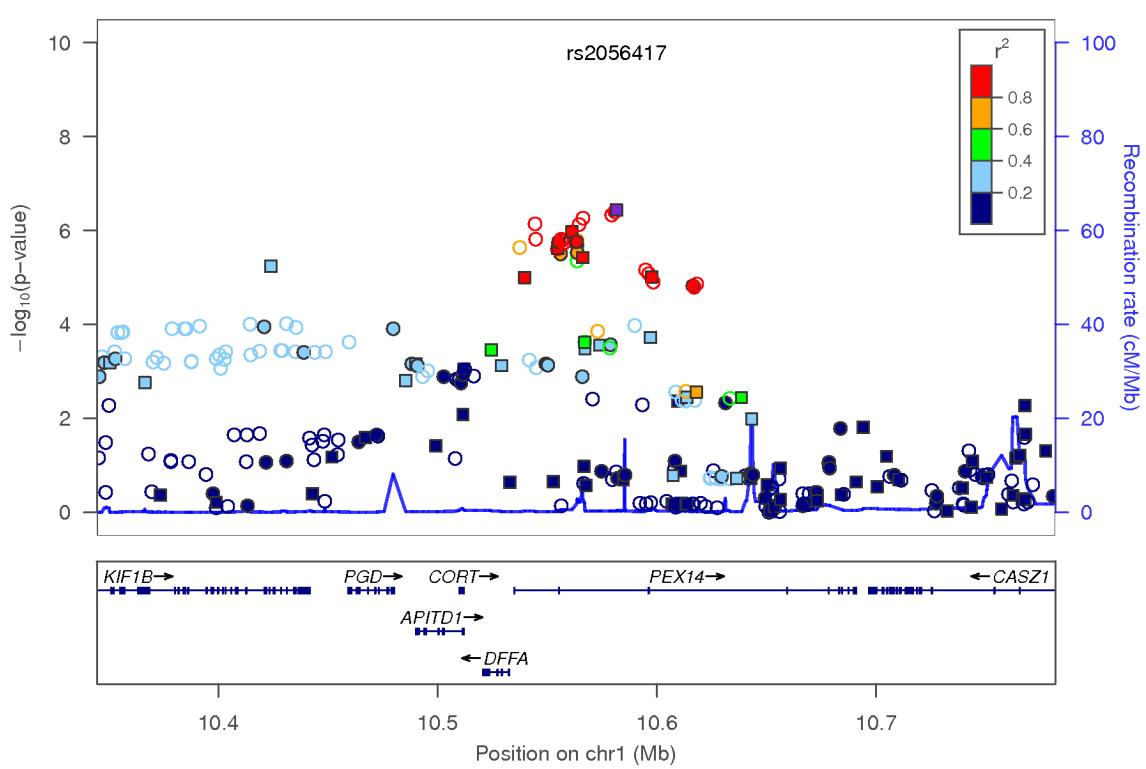
S



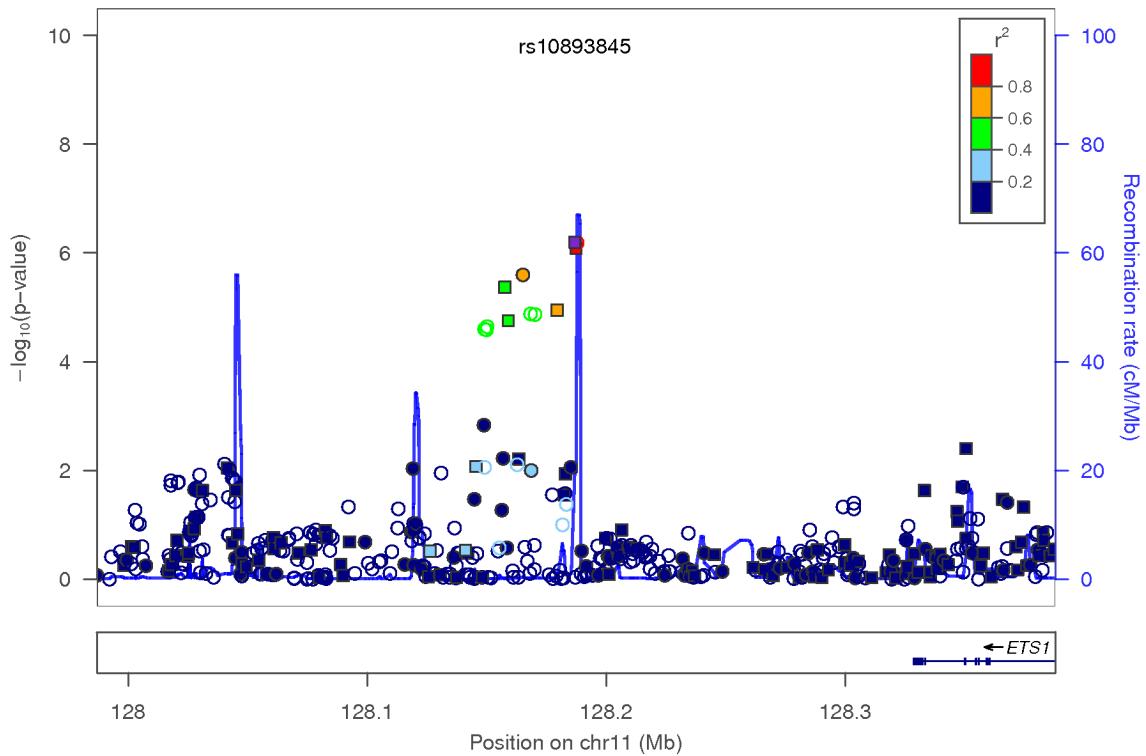
T



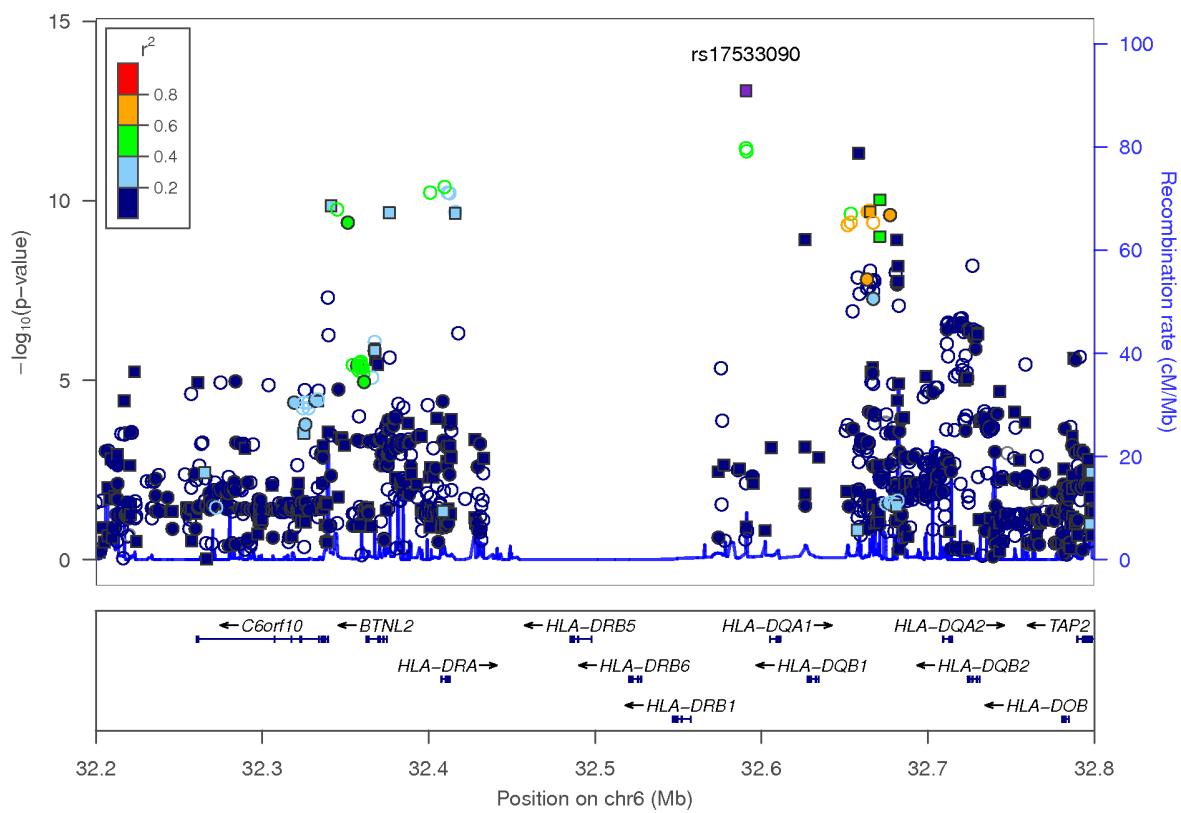
U



V



Supplementary Figure 2. Regional plot of allergy interaction meta-analysis results. Symbol colors indicate linkage disequilibrium with the lead SNP. Open circles indicate imputed variants, filled circles indicate partially genotyped variants, and filled squares indicate fully genotyped variants. Results are in NCBI Build 37 coordinates.



## Supplementary Note

### 23andMe Surveys and Scoring Logic

The following section shows excerpts from the surveys used to define cases and controls for each allergy phenotype in the 23andMe cohort.

#### ***Survey: Allergies and Asthma***

##### [Q1] Have you had an allergic reaction to any of the following items?

	Yes	No	I'm not sure
Foods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q1a] Plants (including pollen)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q1b] Animals (mammals, birds, or insects)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q1c] Dust mites	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Molds	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Latex	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Medicines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vaccines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Something else	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

##### [Q2: conditional on Q1a] Have you had an *allergic reaction* to any of the following types of plants?

	Yes	No	I'm not sure
[Q2a] Grasses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q2b] Trees	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q2c] Weeds	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Another plant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

##### [Q3: conditional on Q1a] Have you had an allergy test administered by a medical professional which indicated that you are allergic to the following types of plants?

	Yes	No	I'm not sure
[Q3a] Grasses	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q3b] Trees	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q3c] Weeds	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Another plant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**[Q4] Have you had an *allergic reaction* to any of the following types of animals?**

	Yes	No	I'm not sure
[Q4a] Cats	●	●	●
Dogs	●	●	●
Honeybees	●	●	●
Red ants	●	●	●
Yellow jackets	●	●	●
Wasps	●	●	●
Another animal	●	●	●

**[Q5: conditional on Q4] Have you had an allergy test administered by a medical professional which indicated that you are allergic to the following types of animals?**

	Yes	No	I'm not sure
[Q5a] Cats	●	●	●
Dogs	●	●	●
Honeybees	●	●	●
Red ants	●	●	●
Yellow jackets	●	●	●
Wasps	●	●	●
Another animal	●	●	●

**[Q6: conditional on Q1c] Have you had an allergy test administered by a medical professional which indicated that you are allergic to the following?**

	Yes	No	I'm not sure
Dust mites	●	●	●
Molds	●	●	●
Latex	●	●	●
Medicines	●	●	●
Vaccines	●	●	●
Something else	●	●	●

**[Q7: asked separately for each allergen conditional on Q1, Q2, and Q4] What type of reaction did you have after being exposed to [allergen]? Please check all that apply.**

- Abdominal pain or vomiting
- Diarrhea
- [Q7a] Difficulty swallowing or difficulty speaking
- Drop of blood pressure or passing out
- [Q7b] Hives (red, itchy, or swollen skin)
- [Q7c] Itching in your mouth
- [Q7d] Itchy or runny nose
- Nausea
- [Q7e] Wheezing or asthma
- None of the above

### ***Survey: Your Medical History***

**[Q8] Have you ever been diagnosed by a doctor with one of the following types of allergies?**

	Yes	No	I'm not sure
[Q8a] Seasonal allergies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Asthma	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
[Q8b] Environmental (but not seasonal) allergy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Food allergy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Drug allergy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### ***Survey: Asthma***

**[Q9] Do you currently have allergies?**

- Yes
- No
- I'm not sure

## **Logic for Selecting 23andMe Cases and Controls**

### *Pollen allergy*

- Cases were defined by a 'Yes' answer to [Q1a] and one or more of [Q2a], [Q2b] or [Q2c], as well as either a 'Yes' answer to the corresponding [Q3a], Q3b] or [Q3c], or one or more of [Q7a], [Q7b], [Q7c], [Q7d], or [Q7e].
- Controls were defined by a 'No' answer to [Q1a], or to all of [Q2a], [Q2b], and [Q2c].

### *Cat allergy*

- Cases were defined by a 'Yes' answer to [Q1b] and [Q4a], as well as either a 'Yes' answer to either [Q5a], or one or more of [Q7a], [Q7b], [Q7c], [Q7d], or [Q7e].
- Controls were defined by a 'No' answer to [Q1a], or to all of [Q2a], [Q2b], and [Q2c].

### *Dust mite allergy*

- Cases were defined by a 'Yes' answer to [Q1c], as well as either a 'Yes' answer to [Q6], or one or more of [Q7a], [Q7b], [Q7c], [Q7d], or [Q7e].
- Controls were defined by a 'No' answer to [Q1c].

Additional shared controls for all three allergy phenotypes were then identified, who (1) did not qualify as cases for any allergy phenotype; (2) either responded 'No' to both of [Q8a] and [Q8b], or to [Q9]; and (3) had not responded 'Yes' to any of these questions.

Finally, individuals who were classified as a case for at least one allergy phenotype were classified as controls for other allergy phenotypes for which they were not already classified as cases or controls.

## **Allergy Symptom Analysis**

In this analysis, we reclassified allergy cases and controls based on their responses to symptom question [Q7]. We defined allergic contact dermatitis cases by a positive response to [Q7b]; allergic rhinitis cases by a positive response to [Q7d]; and allergic asthma cases by a positive response to [Q7e], for any allergy type. Controls gave only negative responses for the corresponding symptom.

## Suggestive Associations with Allergy

Two of our suggestive associations, in the 8q21.13 and 2p25.1 regions, were strongly replicated in the companion study (Table 4). In the 8q21.13 region, our index SNP rs13275219 falls between *TPD52*, or tumor protein D52, and *ZBTB10*, or zinc finger and BTB domain containing 10. An association in this region (rs7000782,  $r^2=0.51$ ) was tentatively identified in a discovery GWAS for eczema<sup>14</sup> but failed to replicate in that study. In the 2p25.1 region, our index SNP rs10174949 is 350 kb upstream of *ID2*, or inhibitor of DNA binding protein 2, a transcription factor with roles in the differentiation of lymphoid tissue inducer cells, natural killer cells, and other innate lymphoid cell lineages<sup>74</sup>. In a mouse model, *ID2* knockouts show substantially reduced allergic symptoms in a model of allergic rhinitis<sup>75</sup>.

In the 16p13.13 region, variants in strong LD with our index SNP rs725613 have been widely associated with autoimmune disease in genome-wide association studies, including type 1 diabetes<sup>76</sup>, multiple sclerosis<sup>77</sup>, celiac disease<sup>33</sup> and primary biliary cirrhosis<sup>78</sup>. A recent report suggests that the variants identified in these studies modify expression of the nearby *DEXI* gene<sup>79</sup>, which is differentially expressed in lung epithelia in emphysema<sup>80</sup>. An association has also been reported with allergic rhinitis but did not reach genome-wide significance<sup>17</sup>.

Among the remaining suggestive associations, in the 16p12.1 region, index SNP rs2107357 falls between T helper cytokine receptor genes *IL4R* and *IL21R*. IL-4 and IL-21 cytokines are associated with differentiation of Th2 and Th17 cell lineages, respectively<sup>81,82</sup>. In the 11q24.3 region, index SNP rs10893845 is 150 kb downstream of the *ETS1* gene. *ETS1* is a transcription factor known to contribute to Th2 cytokine regulation<sup>83</sup>, and inhibits differentiation of Th17 cells<sup>84</sup>. Variation in or near *ETS1* has been associated with autoimmune diseases including celiac disease<sup>33</sup>, rheumatoid arthritis<sup>85</sup>, and systemic lupus erythematosus<sup>86</sup> but these variants are not in strong LD with our index SNP. In the 1p36.22 region, our association with rs2056417 in *PEX14*, or peroxisomal biogenesis factor 14, does not have an obvious connection with immune function.

## Allergy Loci Previously Associated with Asthma

The 5q22.1 region spanning index SNP rs1438673 includes the *TSLP* (Thymic Stromal Lymphopoietin) gene, which has been described as a master switch for allergic inflammation<sup>87</sup> responsible for promoting differentiation of pro-inflammatory Th2 cells<sup>88</sup>. In humans, variants in this region are strongly associated with asthma across multiple ethnic groups<sup>9,10,89</sup>; while our index SNP is in low LD with rs1837253 reported in these studies ( $r^2=0.05$ ), we also see good evidence of association with rs1837253 (meta-analysis  $P=2.3\times 10^{-10}$ ). We see evidence for a secondary peak near SLC25A46 (rs12659961: meta-analysis  $P=1.9\times 10^{-9}$ ), which is not in LD with rs1438673 ( $r^2<0.01$ ), but is in moderate LD with rs17513503 ( $r^2=0.3$ ), reported to be associated with grass sensitization<sup>17</sup>.

In the 11q13.5 region spanning index SNP rs2155219, *LRRC32* (also known as GARP), encodes a protein that mediates the suppressive effects of regulatory T (Treg) cells<sup>90</sup>. Regulatory T cells help suppress responses of Th2 cells to maintain functional tolerance. Patients with IPEX syndrome, which interferes with development of the Treg lineage, develop exaggerated IgE responses and atopic dermatitis<sup>91</sup>. This region has been associated with atopic dermatitis<sup>12</sup>, asthma<sup>11</sup>, allergic rhinitis<sup>17</sup>, hay fever<sup>92</sup>, ulcerative colitis<sup>93</sup> and Crohn's disease<sup>94</sup>.

The 2q12.1 region around rs10189629 includes a cluster of cytokine genes. The *IL1RL1* gene encodes the ST2 receptor, which binds to IL-33, and is thought to be involved in stimulating production of Th2 cell cytokines such as IL-13 that ultimately promote allergic inflammation<sup>95</sup>. SNPs in this region are also significantly associated with both atopic and autoimmune phenotypes, including celiac disease<sup>25,26</sup>, plasma eosinophil count<sup>96</sup>, asthma<sup>8,10</sup> and Crohn's disease<sup>27</sup>.

Near index SNP rs7032572 in the 9p24.1 region, *IL33* encodes a cytokine required for differentiation of Th2 cells, and is an important activator of Th2 and mast cells in response to inflammation<sup>97</sup>. Its receptor, *IL1RL1*, is also a key asthma gene. Variation near *IL33* (rs1342326,  $r^2=0.95$ ) has been previously associated with asthma<sup>8</sup>.

In the 17q12 region near index SNP rs9303280, GSDMA (gasdermin) is known to induce apoptosis in gastric pit cells in the stomach<sup>98</sup>, but has also been associated with asthma and

total number of white blood cells in genome-wide association studies<sup>8,99</sup>. *GSDMA* is also about 40kb away from *ORMDL3*, a gene that has been implicated strongly in asthma<sup>100</sup> and with a number of autoimmune diseases<sup>91,101,102</sup>. The *ORMDL3* gene product may be involved in intracellular Ca<sup>2+</sup> signaling<sup>103</sup>. Our index SNP is in moderately strong LD ( $r^2>0.5$ ) with the variants previously reported to be associated with asthma in this region<sup>8,10,100</sup>.

The role of HLA class II genes is to present mainly exogenous peptides primarily to helper T cells, so it is unsurprising that they have frequently been associated with different allergic responses<sup>104</sup>. SNPs in this region have been identified in a number of GWAS of immune-system related disorders, including asthma<sup>8,9</sup>, celiac disease<sup>33</sup>, vitiligo<sup>34</sup>, ulcerative colitis<sup>105</sup>, systemic sclerosis<sup>106</sup>, multiple sclerosis<sup>107</sup>, rheumatoid arthritis<sup>108</sup>, systemic lupus erythematosus<sup>109</sup>, and selective IgA deficiency<sup>110</sup>. Our index SNP rs6906021 is in moderate LD ( $r^2=0.43$ ) with rs9273349, the lead SNP in the asthma studies<sup>8,9</sup>.

The *SMAD3* gene encodes a key mediator of transforming growth factor beta (TGF- $\beta$ ) signaling. Variation in *SMAD3* has previously been associated with asthma<sup>8</sup> and with Crohn's disease<sup>31</sup>. A mouse *SMAD3* knockout leads to an exaggerated asthma phenotype<sup>111</sup>, as well as altered skin sensitization<sup>112</sup>. Our index SNP rs17228058 is only moderate LD ( $r^2=0.3$ ) with asthma association rs744910, but is in nearly complete LD ( $r^2=0.99$ ) with the Crohn's association.

Index SNP rs962993 in the 10p14 region is in a gene poor region about 1 mb downstream of *GATA3*, or GATA-binding protein 3. *GATA3* was originally identified as a T cell specific transcription factor and is thought to be a master regulator of Th2 cell differentiation<sup>113,114</sup>. This region has been associated with asthma in a Japanese cohort<sup>9</sup>; our rs962993 is about 80 kb from the reported SNP, rs10508372, but the two are not in linkage disequilibrium in either Europeans or Asians ( $r^2 < 0.01$ ).

## **Supplemental References**

73. Pruim, R.J. et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-2337 (2010).
74. Rankin, L. & Belz, G.T. Diverse roles of inhibitor of differentiation 2 in adaptive immunity. *Clin. Dev. Immunol.* **2011**, 281569 (2011).
75. Kim, D.-Y. et al. Implications of nasopharynx-associated lymphoid tissue (NALT) in the development of allergic responses in an allergic rhinitis mouse model. *Allergy* **67**, 502-509 (2012).
76. Burton, P.R. et al. Genome-wide association study of 14000 cases of seven common diseases and 3000 shared controls. *Nature* **447**, 661-678 (2007).
77. Hafler, D.A. et al. Risk alleles for multiple sclerosis identified by a genomewide study. *N. Engl. J. Med.* **357**, 851-862 (2007).
78. Mells, G.F. et al. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat. Genet.* **43**, 329-332 (2011).
79. Davison, L.J. et al. Long-range DNA looping and gene expression analyses identify DEXI as an autoimmune disease candidate gene. *Hum. Mol. Genet.* doi:10.1093/hmg/ddr468 (2011).
80. Edgar, A.J., Birks, E.J., Yacoub, M.H. and Polak, J.M. Cloning of dexamethasone-induced transcript: a novel glucocorticoid-induced gene that is upregulated in emphysema. *Am. J. Respir. Cell. Mol. Biol.* **25**, 119-124 (2001).
81. Swain, S.L. et al. IL-4 directs the development of Th2-like helper effectors. *J. Immunol.* **145**, 3796-3806 (1990).
82. Korn, T. et al. IL-21 initiates an alternative pathway to induce proinflammatory Th17 cells. *Nature* **448**, 484-487 (2007).
83. Strempel, J.M., Grenningloh, R., Ho, I.C. & Vercelli, D. Phylogenetic and functional analysis identifies Ets-1 as a novel regulator of the Th2 cytokine gene locus. *J. Immunol.* **184**, 1309-1316 (2010).

84. Moisan, J., Grenningloh, R., Bettelli, E., Oukka, M. & Ho, I.C. Ets-1 is a negative regulator of Th17 differentiation. *J. Exp. Med.* **204**, 2825-2835 (2007).
85. Ocada, Y. et al. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat. Genet.* **44**, 511-516 (2012).
86. Yang, W. et al. Genome-wide association study in Asian populations identifies variants in ETS-1 and WDFY4 associated with systematic lupus erythematosus. *PLoS Genet.* **12**, e1000841 (2010).
87. Liu, Y.J. Thymic stromal lymphopoitin: master switch for allergic inflammation. *J. Exp. Med.* **203**, 169–273 (2006).
88. Hammad, H. & Lambrecht, B.N. Dendritic cells and epithelial cells: linking innate and adaptive immunity in asthma. *Nat. Rev. Immunol.* **8**, 193–204 (2008).
89. He, J.Q. et al. A thymic stromal lymphopoitin gene variant is associated with asthma and airway hyperresponsiveness. *J. Allergy Clin. Immunol.* **124**, 222-229 (2009).
90. Wang, R., Wan, Q., Kozhaya, L., Fulii, H. & Unutmaz, D. Identification of a regulatory T cell specific cell surface molecule that mediates suppressive signals and induces Foxp3 expression. *PLoS One* **3**, e2705 (2008).
91. Larche, M. Regulatory T cells in allergy and asthma. *Chest* **132**, 1007–1014 (2007).
92. Marenholz, I. et al. The eczema risk variant on chromosome 11q13 (rs7927894) in the population-based ALSPAC cohort: a novel susceptibility factor for asthma and hay fever. *Hum. Mol. Genet.* **20**, 2443-2449 (2011).
93. Anderson, C.A. et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**, 246–252 (2011).
94. Barrett, J.C. et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* **40**, 955–962 (2008).
95. Hayakawa, H., Hayakawa, M., Kume, A. & Tominaga, S. Soluble ST2 blocks interleukin-33 signaling in allergic airway inflammation. *J. Biol. Chem.* **282**, 26369–26380 (2007).

96. Gudbjartsson, D.F. et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet.* **41**, 342–347 (2009).
97. Schmitz, J. et al. IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokines. *Immunity* **23**, 479-490 (2005).
98. Saeki, N. et al. GASDERMIN, suppressed frequently in gastric cancer, is a target of LMO1 in TGF-beta-dependent apoptotic signalling. *Oncogene* **26**, 6488–6498 (2007).
99. Soranzo, N. et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat. Genet.* **41**, 1182–1190 (2009).
100. Moffatt, M.F. et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007).
101. Barrett, J.C. et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* **41**, 703–707 (2009).
102. Liu, X. et al. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat. Genet.* **42**, 658–660 (2010).
103. Cantero-Recasens, G., Fandos, C., Rubio-Moscardo, F., Valverde, M.A. & Vicente, R. The asthma-associated ORMDL3 gene product regulates endoplasmic reticulum-mediated calcium signalling and cellular stress. *Hum. Mol. Genet.* **19**, 111–121 (2010).
104. Vercelli, D. Discovering susceptibility genes for asthma and allergy. *Nat. Rev. Immunol.* **8**, 169-182 (2008).
105. Silverberg, M.S. et al. Ulcerative colitis loci on chromosomes 1p36 and 12q15 identified by genome-wide association study. *Nat. Genet.* **41**, 216–220 (2009).
106. Radstake, T.R. et al. Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nat. Genet.* **42**, 426–429 (2010).
107. Sanna, S. et al. Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nat. Genet.* **42**, 495–497 (2010).

108. Plenge, R.M. et al. TRAF1-C5 as a risk locus for rheumatoid arthritis - a genomewide study. *N. Engl. J. Med.* **357**, 1199–1209 (2007).
109. Hom, G. et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N. Engl. J. Med.* **358**, 900–909 (2008).
110. Ferreira, R.C. et al. Association of IFIH1 and other autoimmunity risk alleles with selective IgA deficiency. *Nat. Genet.* **42**, 777–780 (2010).
111. Anthoni, M. et al. Smad3-signalling and Th2 cytokines in normal mouse airways and in a mouse model of asthma. *Int. J. Biol. Sci.* **3**, 477-485 (2007).
112. Anthoni, M., Fyhrquist-Vanni, N., Wolff, H., Alenius, H. & Lauerma, A. Transforming growth factor- $\beta$ /Smad3 signalling regulates inflammatory responses in a murine model of contact hypersensitivity. *Br. J. Dermatol.* **159**, 546-554 (2008).
113. Mowen, K.A. & Glimcher, L.H. Signaling pathways in Th2 development. *Annu. Rev. Immunol.* **21**, 713-758 (2004).
114. Ho, I.C. & Pai, S-Y. GATA-3: not just for Th2 cells anymore. *Cell. Mol. Immunol.* **4**, 15-29 (2007).